UNIVERSITÀ DEGLI STUDI DI CATANIA
FACOLTÀ DI LETTERE E FILOSOFIA

DIPARTIMENTO DI SCIENZE DELLA CULTURA, DELL'UOMO E DEL TERRITORIO

DOTTORATO DI RICERCA IN STORIA
(STORIA DELLA CULTURA, DELLA SOCIETÀ E DEL TERRITORIO IN ETÀ MODERNA)
XXIV CICLO

MONICA SASO

STRUMENTI INFORMATICI, DIGITALIZZAZIONE E MULTIMEDIALITÀ PER LE FONTI STORICHE E LA RICERCA NELL'AREA UMANISTICA

TESI DI DOTTORATO DI RICERCA

Coordinatore:
Chiar.mo Prof. Enrico Iachello

Tutor:
Chiar.mo Prof. Enrico Iachello

TRIENNIO ACCADEMICO 2008 - 2011

Indice

Introduzione

La rivoluzione digitale, la ricerca

La rivoluzione digitale è oggi una realtà. Le definizioni e le interpretazioni sono varie e molteplici, ma tutte d'accordo nell'affermare che questo fenomeno sia tangibile. Nel 1995, Nicholas Negroponte nel suo *Being Digital*[1] prefigurava la rivoluzione digitale nell'avvento dei bit sugli atomi, prevedendo come i flussi di dati digitali sarebbero stati in alla base dell'interazione umana, facendo del computer lo strumento indispensabile dell'interazione stessa. Una teoria visionaria per quel periodo, negli anni in cui internet era ancora nella culla. La rivoluzione digitale per Negroponte doveva essere legata a tre fattori: la connessione di tutti i computer ad un'unica rete, l'aumento della velocità di elaborazione dei dati a livello esponenziale e la convergenza digitale nella multimedialità. Tre elementi che oggi sono pienamente soddisfatti, forse superati. In meno di quindici anni, il mondo digitale previsto da Negroponte è di fatto una realtà.

Forse il dato più evidente di questo processo di trasformazione è stata la diffusione, probabilmente mai prima osservata con una simile rapidità, con cui i prodotti digitali si sono

[1]Nicholas Negroponte, *Essere digitali*, Sperling & Kupfer, Milano, 2004.

affermati. Ma non sono in sé gli *oggetti* digitali ad aver fatto la rivoluzione. Si tratta di analizzare in che modo i prodotti digitali abbiano modificato l'accesso all'informazione. Tutto si gioca sulla disponibilità di accesso ai dati.

L'impatto che la rivoluzione digitale ha avuto non è dovuta al cambiamento tecnologico, al mondo scientifico. O meglio, lo è, ma solo in parte. La digitalizzazione e il mondo digitale hanno prodotto soprattutto cambiamenti sociali, economici, politici. È una rivoluzione che ha toccato tutti i punti d'incontro tra ogni individuo e il modo in cui si approccia a qualunque atto comunicativo. Non certo perchè oggi i computer, gli smartphone, internet traducono le nostre conoscenze in linguaggio binario. La *convergenza al digitale* – quel fenomeno per cui diverse tipologie di informazione tradizionalmente legate a mezzi di comunicazione diversi (la stampa, il telefono, la televisione e via dicendo) – in questo caso è il *fenomeno* e non il *cambiamento*. Facciamo un esempio. Uno studioso italiano consulta un testo che si trova in una biblioteca australiana, con il suo *ipad* collegato a internet, mentre viaggia in treno. La rivoluzione non sta nel fatto che possa leggere un testo che è composto da stringhe di codice binario. La rivoluzione sta nel fatto che è cambiato il suo modo di fare ricerca. È cambiato il suo modo di accedere alle informazioni. È cambiato il suo modo di *conoscere*.

Fino a qualche anno fa, avremmo parlato di rivoluzione digitale come di quel fenomeno che avrebbe permesso di codificare

in forma di dato informatico ogni tipologia e contenuto di informazione, diffondendolo a un numero potenzialmente infinito di persone per mezzo delle reti di comunicazione elettronica. Abbiamo già superato questo "gradino" e bisogna studiare e comprendere quali saranno i possibili futuri sviluppi.

Il *paradigma* del digitale

La rivoluzione digitale, si è detto, è quindi principalmente dovuta all'impatto che il digitale ha avuto sulla società, sulla cultura, sulla comunicazione tra la fine del millennio scorso e l'inizio del nuovo. E se è vero che ha costituito una vera rivoluzione, deve aver prodotto un cambiamento di *paradigma,* come definito da Thomas Kuhn ne *La struttura delle rivoluzioni scientifiche*[2]: un fenomeno paragonabile alla diffusione della stampa di Gutenberg, della macchina a vapore, della polvere da sparo. Con una differenza, una caratteristica originale: il livello di diffusione *globale*. Paolo Ferri ne *La Rivoluzione digitale* sostiene che

> per la prima volta su scala mondiale la sfera economica, il mondo della produzione, e quella del sapere e della conoscenza sembrano convergere integralmente[3].

[2] Thomas Kuhn, *La struttura delle rivoluzioni scientifiche*, Einaudi, Torino, 1999.
[3] Paolo Ferri, *La rivoluzione digitale*, Mimesis, Milano, 1999.

È in parte quello che Nicholas Negroponte sostiene in uno dei suoi tre criteri cardine secondo i quali si fonda la sua interpretazione di rivoluzione digitale. Negroponte sostiene infatti che la trasformazione in atto non sia fine a sé stessa, non sia esclusivamente scientifica o tecnologica, ma che porti con sé una trasformazione economica e sociale, con una ulteriore precisazione, che è ancora una differenza rispetto a tutte le rivoluzioni del passato: il passaggio dalla società di massa a quella dei *personal media*.

La rivoluzione digitale offre una straordinaria opportunità di condivisione della conoscenza, informazione e crescita culturale. Ha cambiato il modo di fare ricerca, non solo storica; è ormai un obiettivo essenziale per garantire la conservazione degli oggetti culturali, nel presente e per le generazioni future. La ricerca e la fruizione sono potenziate grazie all'integrazione dei diversi media, alla comunicazione e alla diffusione attraverso internet, alle potenzialità delle *cyberinfrastrutture*, che garantiscono e velocità d'accesso.

La possibilità sempre più estesa di utilizzo delle risorse tecnologiche, informatiche ed elettroniche, ha gettato le basi per una trasformazione radicale e un salto di qualità nella realizzazione di sistemi di automazione e di interconnessione in rete, col fine principale di potenziare la capacità di comunicare cultura.

Se per un verso la tecnologia digitale sembra avere il potere di moltiplicare le possibilità di accesso alla conoscenza, per un altro

è un fattore di forte discontinuità con il passato, considerato da molti come oggetto da difendere o da conservare inalterato. Ma il documento digitale è più che un semplice surrogato informatico di una pagina scritta da conservare nella memoria. Guadagna oggi una fisionomia e una valenza proprie: deve essere efficace in termini di resa virtuale e di comunicazione, e conservare integramente quei "valori" che veicola, assicurando la fedeltà all'originale e quindi garantendo la riproducibilità dei contenuti.

1. Il digitale e la ricerca storica e umanistica

1.1 La rivoluzione digitale: opportunità e sfide

Le tecnologie digitali creano un potente strumento per un nuovo sviluppo della ricerca storica e umanistica, proponendo opportunità e sfide alle discipline informatiche, alle *computing resources* e alle cyberinfrastrutture. La natura ricca ed eterogenea del patrimonio umanistico, insieme alla complessità articolata delle molteplici informazioni connesse mettono alla prova le discipline scientifiche, avvezze a muoversi in campi molto più lineari e dettati da regole e leggi esatte.

La struttura dei dati delle Scienze Umane è intrinsecamente complessa, a differenza di quella delle Scienze Naturali, nelle quali definizioni e campi d'azione sono molto più ristretti. Il dominio naturale della storia e del patrimonio culturale è un insieme caratterizzato da contorni indistinti, governato dalla logica sfumata, in gergo detta *fuzzy*; i contenuti complessi e passibili di molteplici interpretazioni mettono a dura prova le tecnologie della conoscenza, anche le più evolute.

Gli effetti dell'adozione di questi nuovi strumenti sui processi e le pratiche metodologiche pongono la storia e la ricerca al principio di un nuovo percorso di opportunità, anche se ancora poco definito rispetto agli obiettivi e alla direzione di marcia.

Allo stato attuale, i passi compiuti insieme alle tecnologie digitali hanno aiutato a combinare gli sviluppi delle Scienze Informatiche con i classici interessi conoscitivi delle Scienze Umane. Ciò ha apportato notevoli benefici nell'interoperabilità tra le diverse discipline e nelle operazioni che richiedono risorse provenienti da diversi domini. Il rapido aumento del materiale in formato digitale ha sollevato il problema della creazione di adeguati strumenti per la ricerca dei dati, facendo fiorire numerosi progetti sperimentali per lo sviluppo di archivi digitali e *repositories*, insieme al potenziamento dell'accesso online per la fruizione dei contenuti.

Ma la ricerca scientifica e informatica non possono fermarsi di fronte al raggiungimento di questi obiettivi: la creazione, l'accesso, la conservazione e lo scambio di documenti digitali sono soltanto l'inizio di un percorso, che deve essere in grado di investire sempre maggiori risorse per il potenziamento e l'applicazione di tecnologie di calcolo intensivo e metodi di ricerca complessi nell'ambito dell'umanistica.

È soprattutto necessario capire in quale senso l'informatica possa svolgere un ruolo significativo: negli algoritmi e nelle

strutture di dati, nei formalismi e nei linguaggi informatici, negli schemi e nelle forme di modellazione della conoscenza. Dovrebbe essere proprio la storia e gli storici a dover stimolare questo processo, avanzando proposte convincenti per spronare l'informatica alla continua ricerca e produzione di tecniche e metodi nuovi da applicare, insieme a una continua riflessione critica sulle forme digitali di rappresentazione della conoscenza.

1.2 La tradizione umanistica e la nuova cultura informatica

L'introduzione delle tecnologie informatiche nelle singole discipline umanistiche è avvenuta con modi e tempi diversi, grazie a un progresso sempre in crescita nell'adozione di procedure computazionali e di tecniche multimediali. Ciascun ambito di ricerca ha sviluppato metodologie diverse e ha applicato strumenti informatici precisi in relazione alle specificità di ciascuna disciplina.

Gli studi storici hanno ne hanno tratto beneficio grazie alla possibilità di affrontare con uno sguardo nuovo il concetto di fonte, per mezzo dell'inclusione flessibile di media (audio, video, immagini, oggetti tridimensionali) e di una nuova interpretazione dei processi di analisi e sintesi della documentazione raccolta nelle banche dati, con una notevole crescita negli studi di storia

economica e sociale.

Tra gli studi linguistici sono stati affrontati i problemi legati allo studio dei formalismi del linguaggio attraverso la linguistica computazionale e rilevati importanti progressi nelle tecniche di analisi e generazione automatica del linguaggio. Analogamente, nell'ambito degli studi filologici, si è sviluppata una nuova visione sul lavoro di analisi critica e testuale, basata sulla gestione automatica dei manoscritti e dei documenti codificati e memorizzati su supporti digitali.

Gli studi filosofici hanno contribuito a fare dell'informatica stessa un tema di studio, stimolando riflessioni che emergono dall'incontro tra filosofia, scienze cognitive e intelligenza artificiale. Queste considerazioni sono confluite nella filosofia computazionale, un nuovo campo di studi che ha come centro di interesse la conoscenza scientifica e la sua razionalità, influenzando in modo considerevole gli studi più tradizionali di filosofia della mente, del linguaggio e della scienza.

L'informatica applicata all'archeologia ha definito tecniche e metodi nuovi per lo studio topografico, per la gestione e la presentazione dei reperti di scavo, e per la ricostruzione di ambienti del passato, grazie all'uso di tecniche di calcolo e simulazione.

Negli studi letterari, nuovi metodi di indagine e di analisi dei testi basati sul calcolo elaborativo hanno permesso notevoli progressi, ponendo particolare attenzione allo studio del lessico,

all'analisi dello stile e all'attribuzione d'autore anche in forme automatizzate.

Gli studi storico-artistici hanno individuato, infine, nuovi criteri di valutazione e indagine delle opere d'arte, tramite l'analisi numerica delle componenti grafiche, coloristiche e le analisi spettrometriche.

Al di là dei risultati positivi ottenuti da ogni singola disciplina e degli esiti favorevoli del lavoro delle Scienze Informatiche in campi separati per specificità di applicazione, negli ultimi decenni sono stati raggiunti dei progressi, che hanno investito in maniera vantaggiosa il campo umanistico in senso generale. I differenti ambiti di ricerca condividono interessi comuni e trasversali: uno storico, un letterato, e un filosofo avranno in egual modo bisogno di principi, metodologie, strumenti teorici e pratici comuni e generali che provengono dal mondo delle discipline informatiche. In questo senso, il quadro generale dei successi raggiunti è straordinario e molto variegato.

1.3 La storia e il digitale

L'impatto della rivoluzione digitale ha mutato profondamente alcune questioni che sono basilari nel lavoro dello storico, su almeno due versanti diversi: la costruzione del testo storiografico da

una parte, la sua ricezione dall'altra. Anzitutto, il *vecchio* non scompare: come spesso capita nella storia della comunicazione, infatti, i nuovi media convivono con i vecchi e l'avvento dei mezzi digitali non produce la morte di quelli analogici. Archivi e fonti analogiche sono ancora una parte centrale del nostro presente e, con ogni probabilità, lo saranno anche nel futuro: la pubblicazione su carta o su nastro continua e le biblioteche e gli archivi cartacei sono, e continueranno ad essere, rilevanti per l'umanità e gli storici dei media.

1.4 Testi e ipertesti

Numerosi strumenti informatici, ormai largamente diffusi, consentono la realizzazione di edizioni elettroniche di testi, dalle caratteristiche notevolmente diverse rispetto alle tradizionali edizioni cartacee. L'informatica ha cambiato, attraverso l'introduzione degli ipertesti e dell'editoria multimediale, il concetto di testualità, la forma dell'apparato critico, il rapporto fra testo e gli strumenti per la sua analisi e fruizione.

Il testo, nella sua accezione tradizionale, offre un'unica possibilità di lettura. È lineare e consente un percorso unico. Se in questa direzione può risultare il metodo d'approccio più semplice nella presentazione dei contenuti, non consente libertà al lettore-

utente e si muove in maniera totalmente indipendente rispetto al metodo associativo e partecipativo.

L'ipertesto, teorizzato e definito per la prima volta da Theodor Nelson nel 1963, consiste invece in una «scrittura non sequenziale, testo che si dirama e consente al lettore di scegliere; qualcosa che si fruisce al meglio davanti a uno schermo interattivo[4]». A proposito di sequenzialità della lettura e percorso dell'utente all'interno dei testo, Nelson aggiunge:

> L'ipertesto include come a caso particolare la scrittura sequenziale, ed è quindi la forma più generale di scrittura. Non più limitati alla sola sequenza, con un ipertesto possiamo creare nuove forme di scrittura che riflettano la struttura di ciò di cui scriviamo; e i lettori possono scegliere percorsi diversi a seconda delle loro attitudini, o del corso dei loro pensieri, in un modo finora ritenuto impossibile[5].

Nell'accezione di *iper*testo di Nelson, l'accezione di *hyper* si riferisce all'uso che si fa in geometria del termine, che equivale ad essere *esteso* su uno spazio tridimensionale. Un ipertesto è, quindi, un testo a tre dimensioni, un insieme di blocchi o frammenti testuali collegati elettronicamente fra loro secondo una rete di interconnessioni semantiche non sequenziali.

Questa struttura reticolare è segnalata all'utente per mezzo di un'interfaccia che consente al fruitore di identificare gli snodi per spostarsi da un punto all'altro della struttura ipertestuale. I

[4] Theodor Holm Nelson, *Literary Machines* 90.1, Padova, Muzzio, 1992.
[5] *Ibid.*

collegamenti (o *links*) attivi sono parole, frasi, bottoni o immagini cliccabili, agendo sulle quali è possibile aprire o raggiungere un altro frammento testuale. La libertà di azione dell'utente-lettore abbandona la logica lineare del testo tradizionale stampato, a favore di una struttura che si presta a modalità di lettura trasversali ed in cui è impossibile delineare un inizio, una fine, un centro. Ogni fruitore sceglie il percorso, l'ordine e opera una selezione sui contenuti che vuole o non vuole raggiungere.

1.5 Metodologie di rappresentazione e codifica dell'informazione

La necessità di rappresentare in maniera precisa le caratteristiche strutturali e formali delle fonti utilizzate per la ricerca in campo umanistico ha fatto emergere la necessità di associare a esse meta-informazioni descrittive e interpretative, contribuendo allo sviluppo dei linguaggi di marcatura per i documenti. Tale esigenza si è estesa anche alle informazioni non testuali (suoni, immagini, video, dati numerici e geografici), ormai considerate non più semplicemente fatti comunicativi, ma prodotti culturali complessi e strutturati, soggetti a interpretazione e dotati di proprie dimensioni di diretto interesse per lo studioso di Scienze Umane.

1.6 Sviluppo del concetto di documento digitale e strumenti per la circolazione dell'informazione

La rivoluzione digitale delle fonti ha gettato le basi per una profonda riflessione per la definizione di procedure standardizzate e uniformi per la modalità di produzione, gestione, distribuzione e conservazione nel tempo delle risorse digitali. Si è giunti alla precisazione di modalità e strumenti da utilizzare per la produzione dei documenti digitali e per la loro identificazione, manipolazione e salvaguardia nel tempo.

Riviste e pubblicazioni elettroniche, forum di discussione, convegni online, ma anche il semplice uso delle e-mail, hanno profondamente modificato i canali di comunicazione che gli umanisti sfruttano nel proprio lavoro quotidiano, comprese le metodologie della ricerca e la didattica.

1.7 Fonti digitali

Le fonti di natura analogica sono state, sono e continueranno ad essere utilizzate dagli storici.

Le fonti già esistenti, al momento della digitalizzazione, della conversione in codice binario e trasferimento su diverso supporto, divengono *oggetti digitali*. Questo non significa che ne cambi solo la

forma: la leggibilità su carta o a schermo o la possibilità di consultazione anche a distanza. La differenza principale è che si apre un ventagio di possibilità di studio che fino ad oggi non era stato possibile, che pone lo storico di fronte alla possibilità di rileggere questioni fondate con strumenti diversi, più efficienti, con diversi criteri di ricerca, accessibilità e comparabilità dei risultati. Si offre quindi la possibilità di applicare metodi nuovi a vecchi problemi storiografici.

Un testo digitalizzato, non importa di che natura o di che periodo, offre quasi sempre la possibilità di effettuare ricerche di *information retrivial* per mezzo di parole chiave. La consultazione e la ricerca di argomenti di pertinenza diventa centinaia di volte più veloce rispetto alla lettura e allo studio tradizionale, per non parlare della precisione nella selezione dei risultati. In frazioni di secondo si individuano e visualizzano le risposte alla *query*, alla domanda, alla richiesta che poniamo al testo. Se si registra a livello quantitativo una serie di dati su un database, si potrà ottenere automaticamente una comparazione, potenziando la capacità di sintesi e comparazione.

Un'altra questione, forse ancora più importante del cambiamento delle fonti che sono trasformate in oggetti digitali, è quella dell'introduzione di una categoria nuova di fonti, le fonti *born digital*, o*vvero* quelle che nascono già in formato digitale, che

costituiscono oggi le fonti storiche del presente per la futura storiografia.

2. Perché digitalizzare?

La digitalizzazione è un processo di conversione di un segnale o codice analogico in un segnale o codice digitale. Un sistema digitale sfrutta segnali discreti per rappresentare e riprodurre segnali continui sotto forma di numeri. Attraverso il codice binario, un sistema numerico in base due che utilizza due simboli, tipicamente 0 e 1, si è in grado di rappresentare le informazioni in stati del tipo acceso/spento. Il processo di digitalizzazione permette di creare file digitali attraverso la scansione o altre forme di conversione di materiali analogici.

Se, tradizionalmente, lo scopo principale delle biblioteche, degli archivi, dei musei è quello di preservare e diffondere il sapere, oggi questa importante funzione può trarre vantaggio dallo sviluppo tecnologico digitale. La digitalizzazione permette infatti di sottrarre la copia originale ai fenomeni di deterioramento dovuti alla manipolazione del materiale, procurandone una riproduzione in formato digitale, in grado di rappresentare in maniera soddisfacente l'originale, garantendo, quindi, la fruibilità dei contenuti e migliorando le possibilità di accesso ai documenti.

Diversi sono i vantaggi che l'attività di digitalizzazione può offrire: dall'accessibilità da parte di un pubblico vasto a quantità notevoli di documenti rari, alla tutela dei testi originali, alla possibilità di accesso a documenti fisicamente lontani, fino all'opportunità di raccogliere e catalogare i documenti digitalizzati in collane virtuali.

2.1 Le finalità

L'approccio alla digitalizzazione può avere natura e scopi diversi: può essere una pratica di conservazione, volta a proteggere i documenti originali, oppure può essere una pratica di diffusione, indirizzata a migliorare l'accesso ai documenti. In realtà, per sfruttare a pieno le potenzialità dei mezzi informatici, ogni buon progetto di digitalizzazione non dovrebbe propendere né per la prima, né per la seconda finalità, ma procedere parallelamente, valorizzando entrambi gli approcci.

2.1.1 Digitalizzare per conservare

La digitalizzazione intesa come pratica conservativa ha il compito assicurare la salvaguardia dei documenti originali. L'interesse principale è quello di preservare in formato elettronico le opere in modo integrale, sicuro e tale da garantire il futuro a

lungo termine dei materiali. Il *focus* sull'esemplare e sulle sue caratteristiche comporta anche una notevole attenzione verso l'applicazione di una serie di precauzioni, affinché la conversione digitale riproduca la totalità del contenuto informativo.

L'attenzione alle problematiche della conservazione riguarda, una volta digitalizzato, anche l'oggetto digitale, per garantire la longevità e la durevolezza nel tempo dei dati acquisiti

2.1.2 Digitalizzare per diffondere

La digitalizzazione ha spesso finalità di diffusione e divulgazione: la riproduzione dei dati ottenuti ha come scopo la distribuzione degli stessi su scala più o meno ampia. In questa prospettiva, l'accento è posto sulla fruibilità dei contenuti. Il potenziamento della capacità di comunicare informazione grazie all'avvento di internet e dei nuovi media costituisce l'elemento chiave di questo processo, che consente la realizzazione di strumenti reali di accesso e ricerca. La digitalizzazione rappresenta un mezzo importante per garantire un accesso più ampio al materiale culturale e, non va dimenticato, costituisce attualmente l'unico strumento che possa per assicurare una disponibilità sicura di tale materiale alle generazioni a venire.

2.2 Il contesto europeo

2.2.1 La Commissione europea

Il ruolo chiave della digitalizzazione del patrimonio scientifico e culturale è stato riconosciuto nel *Piano d'Azione eEurope 2002* approvato da tutti gli stati membri al Consiglio europeo di Feira nel giugno del 2000.

Tra gli obiettivi principali delle azioni da intraprendere, è individuata la necessità di promuovere la creazione di contenuti culturali europei da collocare sulle reti globali, con lo scopo di favorire la conservazione e la diffusione delle risorse scientifiche e culturali, di stimolare la creazione di nuove opportunità didattiche e di incoraggiare il turismo culturale.

La coordinazione dell'attività dei singoli Paesi è affidata all'ufficio delle Applicazioni al Patrimonio culturale della Direzione Generale Società della Informazione della Commissione, per quanto riguarda le politiche di digitalizzazione applicate al patrimonio culturale.

2.2.2 I principi di Lund

Il 4 aprile 2001, diversi rappresentanti ed esperti della Commissione europea si sono riuniti a Lund, in Svezia, per affrontare le problematiche relative al coordinamento e alla

valorizzazione dei programmi nazionali di digitalizzazione. A conclusione del convegno sono state pubblicate alcune raccomandazioni, definite come *Principi di Lund*[6] con il fine di «mettere a punto un sistema di coordinamento dei programmi di digitalizzazione a livello comunitario», e un *Piano d'azione per il coordinamento dei programmi e delle politiche di digitalizzazione*. Nella premessa ai principi di Lund, si afferma che

> le conoscenze culturali e scientifiche europee formano un patrimonio esclusivo di pubblico valore che riflette la memoria collettiva in perenne evoluzione delle diverse società che caratterizzano l'Europa, fornendo al contempo una solida piattaforma di partenza per lo sviluppo dell'industria europea dei contenuti digitali in una società del sapere a dimensione sostenibile[7].

Il patrimonio europeo deve essere messo a disposizione del pubblico e rientrare in una dinamica di sostenibilità. Per consentire ai cittadini di accedere più facilmente a queste risorse e riuscire a conservare il retaggio culturale del passato, è necessaria la collaborazione degli Stati comunitari per la digitalizzazione dei contenuti europei.

Gli ostacoli maggiori sono stati individuati in in un approccio frammentato delle attività di digitalizzazione, nell'obsolescenza delle

[6] Commissione europea, *Lund Principles*, Svezia, 4 aprile 2001, <http://cordis.europa.eu/ist/digicult/lund-principles.htm>

[7] Versione italiana disponibile all'indirizzo <ftp://ftp.cordis.lu/pub/ist/docs/digicult/lund_principles-it.pdf>.

tecnologie, nella mancanza di standard adeguati e di modalità di accesso semplici e universali a questo patrimonio.

La Commissione europea ha quindi coinvolto i singoli Ministeri responsabili della programmazione e del finanziamento di azioni nazionali degli Stati membri, che hanno istituito un coordinamento stabile la cui finalità è quella di promuovere la nascita di metodologie standard e lo sviluppo delle capacità tecniche essenziali per la creazione di repertori nazionali di contenuti scientifici e culturali.

2.2.3 Il Progetto Minerva

Il progetto Minerva (*MInisterial NEtwoRk for Valorising Activities in digitisation*) ha il compito di realizzare il coordinamento auspicato dalla Commissione europea e dagli Stati membri e di realizzare il Piano d'azione di Lund. Minerva è un progetto finanziato dalla Commissione europea nell'ambito del programma IST (Tecnologie per la Società dell'Informazione) e del V Programma Quadro, con l'obiettivo di creare una rete di Ministeri europei per la cultura, sotto la guida del Ministero per i Beni e le Attività Culturali, per coordinare le politiche nazionali di digitalizzazione.

La strategia di Minerva è quella di diventare un polo di attrazione e di integrazione per istituzioni diverse che operano nel campo della digitalizzazione dei beni culturali. L'allargamento della

rete è considerato sotto diversi profili: stabilire rapporti di collaborazione e divenire un punto di riferimento stabile, anche nei confronti di una serie di altri attori che si muovono nell'ambito dei beni culturali. Tra questi, un'attenzione particolare va alle università e agli istituti titolari di programmi e di attività di ricerca nel campo delle tecnologie applicate ai beni culturali, agli enti e istituzioni pubbliche locali, alle imprese del settore, alle associazioni professionali, e altri progetti coofinanziati dalla Commissione europea o da risorse nazionali che rientrano nel campo d'interesse.

Minerva mira a diventare il principale punto di riferimento della digitalizzazione del patrimonio culturale in Europa, in particolare per quanto riguarda la qualità, le linee guida sulla digitalizzazione, i metadati, la conservazione a lungo termine, l'accessibilità, con l'ausilio di una base di dati multilingue.

Per divulgare i risultati del progetto e renderli disponibili a un vasto pubblico, sono stati pubblicati alcuni lavori, volti a fornire in modo chiaro ed esauriente informazioni specifiche sul panorama europeo della digitalizzazione e sui criteri suggeriti. In quest'ambito si inserisce il manuale *Good practice handbook*[8], una guida alle "buone maniere" nel campo della digitalizzazione del patrimonio culturale.

---

[8] Gruppo di lavoro 6 del Progetto Minerva (Identificazione delle buone pratiche e dei centri di competenza), *Good practice handbook*, versione 1.3, marzo 2004, traduzione italiana a cura di Mario Sebastiani,
<http://www.minervaeurope.org/structure/workinggroups/goodpract/document/buonepratiche1_3.pdf>

2.2.4 Europeana

Europeana (*European Digital Library, Museum, Archive)* è un'iniziativa europea nata per garantire l'accesso a contenuti digitali eterogenei.

Da dicembre 2008 è accessibile online un sito web che dà accesso a circa 2 milioni di oggetti digitali. Le risorse digitali sono selezionate tra quelle già digitalizzate nell'ambito di archivi, biblioteche, musei e archivi audiovisivi.

La biblioteca digitale europea non sarà riunita in un'unica banca dati, ma Europeana fungerà da punto unico d'accesso e collegherà le diverse banche dati sparse per l'Europa. Gli utenti non dovranno più interrogare ogni singola biblioteca digitale per trovare il materiale di cui hanno bisogno. Il punto unico d'accesso, o "portale", sarà gestito dalla Fondazione per la biblioteca digitale europea. Istituita l'8 novembre 2007, la Fondazione riunisce varie istituzioni culturali europee che partecipano insieme al progetto pur mantenendo ciascuno la propria autonomia (ad es. per quanto riguarda le assunzioni e i finanziamenti). Tra i membri fondatori si trovano le associazioni europee di biblioteche, archivi, musei e videoteche. La struttura operativa su cui si appoggia Europeana è ospitata dalla biblioteca nazionale olandese.

2.2.5 BDI - Biblioteca Digitale Italiana

L'idea di dar vita ad un progetto coordinato di Biblioteca Digitale in Italia nasce dalla realizzazione di uno Studio di fattibilità, commissionato dalla Direzione Generale per le Biblioteche, gli Istituti Culturali e il Diritto d'Autore alle società Unysis ed Intersistemi di Roma alla fine del 1999. Lo studio fornisce il quadro di riferimento necessario alla realizzazione di una biblioteca digitale italiana in grado di inserirsi tra le analoghe iniziative europee ed extraeuropee: Biblioteca Digitale Italiana (BDI). Il progetto prevede un Comitato Guida presieduto dal prof. Tullio Gregory e composto da rappresentanti della realtà bibliotecaria statale e regionale, dei musei, dell'università e della ricerca, con il compito di definire il quadro di riferimento culturale e scientifico entro cui collocare le iniziative esistenti, di individuare le priorità degli interventi, di indicare standard e tecnologie comuni da adottare, di raccordare le attività di livello nazionale con le iniziative internazionali. Il Comitato assume come principi operativi due orientamenti fondamentali: i programmi di digitalizzazione, metadati compresi, devono essere definiti centralmente e devono prevedere la scansione di fondi completi e preliminarmente definiti. I documenti digitalizzati devono, inoltre, essere sempre resi disponibili sulla rete internet per permettere la conoscenza e la valorizzazione del patrimonio culturale e scientifico italiano. A questo scopo stabilisce che tutte le risorse digitali devono essere presenti nel Portale

Internet Culturale realizzato nell'ambito del Progetto la Biblioteca Digitale Italiana ed il Network Turistico Culturale (BDI & NTC).

Il progetto BDI si è concentrato sulla digitalizzazione dei cataloghi storici[9] (un primo programma relativo alla scansione, in formato immagine, dei cataloghi storici delle biblioteche pubbliche italiane, più di 200 cataloghi storici, a volume e a schede, di biblioteche italiane appartenenti al Ministero per i Beni e le Attività Culturali, a Enti locali e a Istituti di cultura.), sulla digitalizzazione dei documenti musicali (manoscritti e a stampa che contengono musica notata, con la possibilità di navigare dal record bibliografico verso l'immagine digitalizzata) e digitalizzazione delle pubblicazioni periodiche (scansione di riviste storiche preunitarie, come la Gazzetta di Roma e la Rivista europea, conservate presso varie istituzioni italiane fra cui la Biblioteca Nazionale Centrale di Roma, la biblioteca di Storia Moderna e Contemporanea di Roma e l'Universitaria di Pisa).

3. Progettare la digitalizzazione

3.1 Progettazione e valutazione dei costi

Ogni progetto di digitalizzazione deve essere progettato e pianificato opportunamente. In fase di progettazione devono essere

[9]<http://cataloghistorici.bdi.sbn.it/>

definiti gli obiettivi da perseguire e i risultati attesi. Gli scopi che un'istituzione, sia essa pubblica o privata, si propone di raggiungere possono essere assai diversi: produrre copie d'archivio destinate a preservare la conservazione di alcuni tipi di documenti; trasferire in formato digitale dati già conservati su altri generi di supporti, come i microfilm; integrare i dati su piattaforme digitali preesistenti; distribuire i dati digitali localmente o sul web, sperimentare metodi e tecniche proprie di progetti-pilota.

La definizione precisa degli obiettivi è indispensabile per circoscrivere l'ambito di applicazione del progetto e valutare le modalità e le scelte da compiere per massimizzarne l'efficacia in relazione alle risorse disponibili.

A livello gestionale è innanzitutto necessario individuare le risorse umane e tecnologiche e definire in dettaglio il budget economico e i costi da sostenere. I costi di realizzazione dipendono da molteplici fattori, ma quelli fondamentali sono certamente l'ampiezza e la durata del progetto.

In termini finanziari, gran parte della responsabilità spetta alle caratteristiche fisiche dei documenti da digitalizzare. Ogni tipologia di documenti comporta spese economiche molto diverse in termini di attrezzature specializzate o risorse umane qualificate. La tecnologia digitale comporta, generalmente, ingenti investimenti iniziali e bisogna valutare in maniera appropriata la convenienza in

relazione al profilo qualitativo. Le spese per le attrezzature e i programmi sono tuttavia da considerarsi investimenti a lungo termine, poiché la piattaforma tecnologica acquistata potrà naturalmente essere impiegata per ulteriori lavori o progetti.

Le risorse umane possono essere interne o esterne all'istituzione che si occupa di sviluppare il progetto. Le varie fasi del piano progettuale richiedono competenze diverse e devono essere assegnate a personale specializzato, con una formazione precisa. Le abilità necessarie variano in base alle tecnologie adottate, al trattamento degli originali e alle operazioni di schedatura documentaria. I tecnici informatici dovranno lavorare accanto al personale addetto alla presa e ricollocazione; i dati dovranno inoltre essere revisionati da esperti di biblioteconomia e da responsabili della conservazione o del restauro. Non tutte queste figure sono necessarie, naturalmente, ma è auspicabile formare un *team* quanto più eterogeneo e specializzato possibile.

3.2 Standard tecnici

La scelta di utilizzo di un determinato standard in fase di progettazione permette di favorire la compatibilità e l'interoperabilità per la gestione di un numero limitato di formati e protocolli.

Gli standard possono essere:

de jure riconosciuti di fatto da un'istituzione responsabile della definizione e diffusione di standard, di solito discussi e stabiliti sviluppati attraverso un accordo tra le parti interessate;

de facto standard industriali, largamente condivisi e rispettati dagli utenti senza tuttavia l'esistenza di un riconoscimento formale.

In generale, le procedure connesse all'adozione di standard *de jure* assicurano l'accesso aperto (*open access*) e il libero utilizzo (*open use*), senza costi di implementazione o di proprietà intellettuale (licenze d'uso). Questi standard sono, infatti, concepiti in modo da favorire i bisogni degli utenti piuttosto che gli interessi dei produttori.

La scelta di adottare uno standard garantisce l'accesso diretto ai contenuti da parte degli utenti, senza possedere strumenti specializzati (interoperabilità); una fruibilità e un accesso facilitato nei confronti con di qualsiasi categoria di utente (accessibilità); la longevità delle risorse, mantenendole nella naturale continuità temporale (conservazione a lungo termine) e il rispetto dell'identità dei contenuti e dei diritti di proprietà intellettuale (sicurezza).

Si discuterà dettagliatamente in seguito a proposito dell'utilizzo di standard per i formati, i metadati e altri elementi fondamentali.

3.3 Flusso di lavoro

Il primo passo per dare avvio a un programma di digitalizzazione è quello di definire un flusso di lavoro per stabilire e articolare tutte le fasi e le operazioni di elaborazione. Una programmazione preliminare puntuale delle attività da svolgere permette l'ottimizzazione e il controllo di tutte le attività relative al progetto.

Il documento *Linee guida tecniche per i programmi di creazione di contenuti culturali digitali*[10], stilato dal Ministero per i Beni e le Attività Culturali, propone un flusso di lavoro articolato in nove fasi:

1. catalogazione dell'oggetto da digitalizzare;
2. acquisizione delle immagini digitali, produzione del file con nomenclatura stabilita, produzione dei metadati relativi alle immagini;

[10] Giuliana De Francesco, *Linee guida tecniche per i programmi di creazione di contenuti culturali digitali* (versione italiana 1.8), Ministero per i beni e le attività culturali, Progetto MINERVA WP4 – Gruppo di lavoro italiano "Interoperabilità e servizi" <http://www.minervaeurope.org/structure/workinggroups/servprov/documents/technicalguidelinesita1_8.pdf>.

3. primo controllo qualità sulle immagini;
4. archiviazione dei master digitali delle immagini;
5. produzione dei metadati amministrativi, descrittivi e strutturali;
6. generazione di formati atti alla pubblicazione in rete, secondo una predefinita strategia di fruizione pubblica;
7. pubblicazione nell'area di lavoro interna (intranet) ed esecuzione del secondo controllo di qualità;
8. pubblicazione sul web;
9. collaudo.

Per ottimizzare l'attuazione delle fasi del programma e raggiungere risultati apprezzabili è opportuno documentare e controllare ogni momento di lavoro del progetto di digitalizzazione, dalla fase iniziale alla fase conclusiva, eventualmente con l'ausilio di un *database* gestionale.

Il *workflow* può naturalmente variare in base alle specifiche del progetto da portare a termine, a seconda che ci si trovi ad affrontare un progetto orientato alla pubblicazione in rete, oppure a uso esclusivamente interno all'istituzione o all'azienda.

4. Gli oggetti da digitalizzare

Nella fase di progettazione è necessario individuare e distinguere le tipologie dei beni da digitalizzare, poiché ogni processo varia in base alle caratteristiche fisiche, alle tecniche di digitalizzazione, agli strumenti e alle attrezzature da utilizzare, fino a una diversificazione delle attività di *editing* e *postprocessing* e degli usi ai quali saranno destinate le risorse digitali prodotte. In particolare, elementi quali il formato, le dimensioni, lo stato di conservazione e riproduzione di proprietà come il colore possono significativamente influenzare la scelta delle architetture hardware e software da impiegare.

4.1 Libri semplici

La distinzione tra libri, documenti dattiloscritti e manoscritti è necessaria in funzione all'attenzione da porre sulla ricognizione del testo: nei libri a stampa e nei documenti dattiloscritti l'attenzione è posta sul testo; nelle opere e nei documenti manoscritti l'attenzione è generalmente posta sull'oggetto stesso, cioè sulle sue proprietà grafiche intrinseche.

Alla categoria dei libri semplici appartengono tutte quelle opere per cui si pone un'interesse in modo predominante sul testo. In genere ci si riferisce, quindi, a libri moderni, i cui livelli di usura non sono in genere elevati e per i quali non è necessario porre particolare attenzione allo stato fisico. L'obiettivo principale è infatti quello di migliorare la distribuzione e la visibilità delle opere piuttosto che preservarle e preoccuparsi della loro conservazione.

4.2 Libri manoscritti, libri di architettura, pergamene

A questa categoria appartengono tutte quelle opere in cui l'interesse predominante è ottenere una riproduzione il più possibile fedele al documento originario. La digitalizzazione deve restituire quindi una precisione specifica legata all'immagine più che al testo. I manoscritti, di cui ci si occuperà in modo particolare successivamente, i libri di architettura e le pergamene sono strumenti di indispensabile per la ricerca in campo storico e filologico. Molti studiosi e filologi sono interessati a visionare l'opera originale, sebbene molte opere siano precluse alla consultazione a causa degli alti livelli di usura delle opere originali. L'obiettivo della digitalizzazione è quindi, in questo caso, quella di fornire un surrogato digitale che conservi caratteristiche qualitativamente molto alte per la consultazione, garantendo una

migliore conservazione degli esemplari autentici e impedendo che un'eccessiva manipolazione dei documenti acceleri il naturale processo di invecchiamento dei materiali. Le valutazioni preliminari devono considerare anche altre qualità specifiche dei documenti manoscritti, dei libri di architettura e delle pergamene, quali l'ordinamento, il tipo di legatura, la presenza di cartulazione o di altri elementi notevoli, come quali sigilli, miniature, annotazioni e stemmi.

Il *Gruppo di lavoro per la digitalizzazione del materiale manoscritto*, composto da esperti dell'ICCU, bibliotecari, medievalisti, archivisti e storici dell'arte, si propone l'obiettivo di istituzionalizzare una serie di linee guida per fornire strumenti metodologici e indicazioni tecniche per tutte le fasi del processo di digitalizzazione, comprese le norme relative al contesto giuridico e alla fruizione. Il documento è attualmente in fase di redazione.

4.3 Bandi, manifesti e fogli volanti

A questa categoria appartengono documenti costituiti da uno o più fogli, o da più carte in forma di fascicolo in diversi formati e misure. Si tratta in genere di disposizioni normative emanate da autorità politiche e religiose o produzioni a carattere letterario, politico, religioso, musicale, pubblicitario ecc. Le caratteristiche

funzionali di questi materiali da tenere in considerazione per la digitalizzazione sono la variabilità della dimensione del supporto cartaceo e l'attenzione particolare da prestare alle qualità grafiche dei prodotti: dalla stampa tradizionale a caratteri mobili fino alle moderne forme di *offset*. La digitalizzazione è mirata alla salvaguardia dei supporti originali, spesso prodotti con materiali facilmente deperibili. La finalità è quindi conservativa e molto simile a quella dei manoscritti.

Nel 2005 è stato istituito dall'ICCU e approvato dal Comitato Guida della BDI un *Gruppo di lavoro per la digitalizzazione di bandi, manifesti e fogli volanti*, composto da esperti dell'ICCU, bibliotecari e archivisti, con l'obiettivo di elaborare delle linee guida per fornire strumenti metodologici e indicazioni tecniche condivisibili da biblioteche, archivi e musei.

Il gruppo ha portato avanti nel 2006 uno studio progetti di digitalizzazione di tale materiale completi e ancora in atto. A conclusione dell'attività, sono state stilate le *Linee guida per la digitalizzazione di bandi, manifesti e fogli volanti*, disponibili sia in volume che online.

4.4 Materiale cartografico

Il materiale cartografico rappresenta un patrimonio di grande valore storico, documentario e artistico e comprende beni culturali

preziosi, ma talvolta scarsamente conosciuti, divulgati o valorizzati a fronte delle qualità possedute. Negli ultimi anni sono sorte numerose iniziative, in ambito italiano e internazionale, volte ad integrare l'attività di catalogazione e inventariazione con quella di digitalizzazione facendo uso delle nuove tecnologie. È stata quindi necessaria la discussione per la creazione di uno strumento di coordinamento, promosso dal Comitato Guida della Biblioteca Digitale Italiana. Il Comitato Guida BDI ha costituito il *Gruppo di lavoro per la digitalizzazione del materiale cartografico* al fine di dare unitarietà alle iniziative di digitalizzazione, e consentire la crescita di un patrimonio digitale di qualità e di larga accessibilità, favorendo la consultazione attraverso i nuovi media. Tali obiettivi, secondo le *Linee Guida* redatte dall'ICCU, vanno perseguiti attraverso

> il recupero del patrimonio digitale esistente; la promozione di campagne di inventariazione e catalogazione, collegate ad interventi di digitalizzazione; la programmazione di interventi per la conservazione degli originali e delle risorse digitali prodotte; l'individuazione di criteri e metodologie comuni nelle modalità di digitalizzazione, nella produzione di metadati e nella conservazione a lungo termine degli archivi digitali; la definizione di un set di metadati descrittivi, condivisibile da biblioteche, archivi, musei e soprintendenze[11].

[11] *Linee guida per la digitalizzazione del materiale cartografico*, a cura del Gruppo di lavoro sulla digitalizzazione del materiale cartografico, Roma, ICCU, 2006, pag. 5.

I materiali cartografici posseggono alcune caratteristiche peculiari che devono essere rispettate e prese in considerazione nel processo di digitalizzazione, quali la variabilità delle dimensioni e il rischio di danneggiamento durante la fase di lavorazione. È quindi essenziale esaminare le condizioni e lo stato di conservazione per ciascun esemplare, in modo tale da poter prevedere opportuni interventi di restauro prima della fase di scansione o prevederne un successivo restauro virtuale. Le valutazioni preliminari devono considerare anche altre qualità specifiche dei documenti cartografici, quali l'eterogeneità di formati e dimensioni, la qualità dei colori, le tecniche di realizzazione, la modalità di piegatura, di arrotolamento e di legatura e l'eventuale presenza e tipologia dei contenitori.

4.5 Opere pittoriche

La digitalizzazione di questo tipo di materiali iconografici deve sempre considerare come priorità la fedeltà della riproduzione rispetto al prodotto originale, realizzando una copia digitale di massima qualità visiva. In questo ambito l'importanza dell'originale, generalmente unico, è fondamentale. Un difficoltà può essere rappresentata dalle enormi differenze tra le categorie che afferiscono a questa vasta tipologia di arti visive, sia in termini di tecniche utilizzate (pittura a tempera, a olio, a tempera, acrilica,

acquerello, affresco ecc.) che di supporti fisici utilizzati (tavole, tele, carta, ceramica, vetro, metallo, stoffe, pareti). A ogni modo, la procedura di digitalizzazione è più o meno simile e le differenti dimensioni non influenzano i processi.

Tra gli obiettivi dei progetti di digitalizzazione delle opere pittoriche vi è spesso quello di supportare l'attività di restauro, che comporta l'uso di strumentazioni e accorgimenti ben diversi dalla semplice conservazione di un'immagine e sono necessarie competenze elevate in questo campo.

4.6 Fotografia

I progetti indirizzati alla digitalizzazione di questa particolare categoria devono tener conto di modalità specifiche a seconda di diversi tipi di prodotti: pellicole (6x6, 10x12, 35mm), stampe fotografiche, diapositive e lastre. La procedura di acquisizione richiede l'utilizzo di dispositivi ottimizzati per la resa dei colori ad alta risoluzione.

L'Istituto Centrale per il Catalogo e la Documentazione (ICCD) ha redatto nel 1998 una *Normativa per l'acquisizione digitale delle immagini fotografiche*[12], per regolamentarne le i processi di acquisizione e memorizzazione e i criteri da seguire per la trasmissione delle stesse.

[12] <http://www.iccd.beniculturali.it/Catalogazione/standard-catalografici/aquisizione-digitale-delle-immagini-fotografiche/>

4.7 Audio e musica

Molte biblioteche di conservazione e archivi storici presenti raccolgono beni audio di particolare valore storico e artistico. Attraverso il processo di digitalizzazione è possibile trasferire i contenuti audio registrati su vecchi supporti, come dischi in vinile, audiocassette, nastri magnetici, in formati di largo uso.

Lo scopo principale dei progetti di digitalizzazione dei materiali audio è la conservazione dei documenti originali. Inoltre, una maggiore fruibilità è garantita dalla possibilità di un intervento che migliori la qualità generale del suono.

4.8 Video

Numerosi archivi e biblioteche ospitano nei loro cataloghi una gran mole di materiale multimediale e audiovisivo. La categoria del materiale video riguarda film, documentari, videoregistrazioni (anche di eventi artistici, culturali, musicali, teatrali, sportivi ecc.) e materiali non montati memorizzati su supporti fisici, quali l'anello cinematografico, la bobina di film, la cartuccia di film, la cassetta di film, la videobobina, la videocartuccia, la videocassetta.

La Rai ha messo a punto un piano di recupero, restauro e digitalizzazione di tutto il suo patrimonio di immagini dal gennaio del 1998. Il progetto, denominato *Teche*[13] ha prodotto un catalogo

[13] <http://www.teche.rai.it>

soltanto per uso interno, dal quale vengono estratti e resi disponibili su internet solo piccole anteprime. Anche l'Istituto Luce ha avviato nel 1996 un progetto di digitalizzazione e prevede delle versioni di video in *streaming* accessibili a tutti, e delle versioni di migliore qualità per gli abbonati al servizio e per le biblioteche e i musei.

4.9 Statue e oggetti tridimensionali

Esistono tecnologie attualmente in grado di rilevare una forma fisica e di restituirla in formato digitale attraverso le coordinate 3D. Sono disponibili varie tecniche di digitalizzazione tridimensionale, distinte in base alle finalità e alle tipologia di scansione.

In genere i progetti di digitalizzazione 3D hanno finalità di documentazione e restauro: grazie alla scansione digitale è possibile analizzare l'oggetto, operare delle misurazioni precise e documentare il lavoro svolto, fungendo anche da supporto all'attività di restauro.

Il punto critico, oltre alla necessità di strumentazioni molto costose (scanner laser o simili), sta nel *rendering* finale dell'oggetto. Se infatti è possibile ottenere correttamente la posizione nello spazio dell'insieme dei punti che descrivono la forma dell'oggetto, è ben più complesso raggiungere un risultato fedele all'originale, che rispetti i colori originali e i dettagli.

5. Strumenti per digitalizzare

Sebbene, in genere, la preoccupazione principale a livello di investimento di risorse per un progetto di digitalizzazione sia concentrata sulla scelta dei dispositivi di acquisizione ottica da impiegare per la scansione, è bene ricordare che una postazione di lavoro per la digitalizzazione è composta da un sistema che comprende tre elementi di natura diversa, che devono interagire e dialogare correttamente:

- **Hardware** (computer e monitor)
- **Software** (programmi vari)
- **Periferiche** (dispositivi di acquisizione ottica, dispositivi di storage di dati e stampanti)

Tra gli elementi da prendere in considerazione figurano dunque la potenza di calcolo necessaria (CPU, scheda madre e RAM); il gruppo scheda video/monitor; i supporti da usare per la memorizzazione e i sistemi di trasferimento di dati (hard disk

interni ed esterni, dischi ottici, DAT). Sono poi indispensabili programmi per digitalizzare, elaborare, catalogare le immagini.

Dopo questa preliminare classificazione, si può procedere ad analizzare tutte le caratteristiche tecniche e di funzionamento degli strumenti principali per l'acquisizione digitale dei documenti: gli scanner.

5.1 Lo scanner

Lo scanner è una periferica in grado di acquisire in modalità ottica un oggetto fisico. Attraverso una serie di componenti meccaniche e ottiche, l'immagine è acquisita, interpretata come insieme di punti e, grazie ad un software interno, convertita in bytes per essere inviata al computer. L'immagine è restituita in formato digitale, che può essere memorizzato e rielaborato. Il processo che viene attuato è definito *campionamento*.

5.2 Campionamento dell'immagine

Quando si campiona un'immagine la si traduce in un insieme di singoli punti, detti pixel, ciascuno dei quali è definito da valori numerici. Ciascun valore rappresenta un colore e l'insieme di tutti i punti è in grado di rappresentare più o meno fedelmente

l'immagine originale. Semplificando, il processo è simile a quello della composizione di un mosaico: l'immagine risulta formata da numero elevatissimo di singole caselline quadrate, ciascuna delle quali rappresenta un colore unico. La somma complessiva produce un'immagine completa e, se il numero delle caselle è sufficientemente alto, l'occhio non riesce più a percepire la presenza dei punti in maniera discreta.

Durante il processo di acquisizione, infatti, è applicata sull'immagine una griglia quanto più fine possibile, che la suddivide in un reticolato; il valore di finezza e precisione della griglia (o *raster*) permette di rappresentarla in modo dettagliato e di ottenere una scansione di qualità.

L'immagine che l'occhio umano percepisce è un insieme infinito di punti luminosi, che sono avvertiti senza soluzione di continuità, a rappresentare un quadro visivo unico. Il computer deve sempre avere a che fare con insiemi finiti di informazioni di un unico tipo: successioni di 0 e 1, ossiadi "interruttori" accesi e spenti.

Il campionamento è processato da due dispositivi: un gruppo optoelettronico che rileva i colori presenti sul documento in scansione e li traduce in segnali elettromagnetici e un convertitore analogico-digitale che associa a ciascun tipo di segnale un valore numerico.

5.3 Requisiti e caratteristiche

La scelta di uno scanner è vincolata da una serie di esigenze. È necessario conoscere i requisiti fondamentali e le caratteristiche generali per valutare l'acquisto di una particolare tipologia e del relativo modello. Questi, infatti, possono produrre risultati molto diversi in base ai tipi di documenti da acquisire, le dimensioni, i tipi di supporto e i tempi da rispettare per portare a compimento il lavoro.

Tra i requisiti da considerare, quelli fondamentali sono la tecnologia del sensore, la risoluzione, la profondità di colore, la gamma di densità, il formato e la velocità di acquisizione.

5.3.1 Tecnologia del sensore

Un sensore di immagini, o sensore ottico, è un dispositivo sensibile alla luce.che converte una immagine ottica in un segnale elettrico. Il sensore è sempre accoppiato a un convertitore analogico-digitale (*A/D converter*), che trasforma l'informazione in dato digitale. Nell'industria elettronica attuale vengono utilizzati diversi tipi di sensori.

Sensore PMT (*photomultiplier tubes*). Il sistema di scansione con tubi fotomoltiplicatori PMT è eseguito da tre fotomoltiplicatori RGB. La sorgente luminosa emessa è una luce alogena al tungsteno,

il cui fascio viene concentrato con lenti e fibre ottiche in modo da illuminare una porzione molto piccola dell'oggetto. La luce riflessa dall'oggetto è raccolta da piccoli specchi semitrasparenti e inclinati, che rimandano una piccola quantità di luce ad altri specchi. Quindi, la luce passa attraverso un appropriato filtro colorato e diretto al corrispondente fotomoltiplicatore, dove avviene il processo di amplificazione ottica. Gli elettroni emessi quando la luce colpisce il catodo del fotomoltiplicatore viaggiano attraverso strati di diodi, che a loro volta emettono ulteriori elettroni, amplificandoli fino al punto di convertire la luce in segnali elettrici. L'anodo del fotomoltiplicatore misura le variazioni di questi segnali, che vanno ai convertitori per essere registrati in segnali digitali.

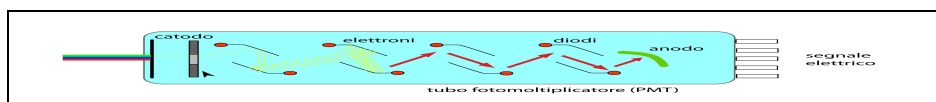



Figura 5.1: schema di un sensore PMT

Sensori CCD (*charged couple devices*)/ CMOS (*complementary metal oxide semiconductor*). Si tratta di sensori che reagiscono alla luce e si avvalgono di tre passaggi di rivelazione, uno per ciascun colore RGB. La luce emessa da una lampada è riflessa dall'originale, poi da uno specchio che la indirizza su una serie di lenti che la focalizzano sul sensore CCD. Qui è infine trasformata in impulsi elettrici, a loro volta trasformati in dati digitali tramite un convertitore analogico digitale.

Il CCD è costruito in matrici di celle unitarie (pixel) che possono trasferire la carica elettrica le une con le altre. Facendo scorrere prima in senso verticale e poi in senso orizzontale le cariche tra le celle, è possibile serializzare l'informazione, così come avviene in tutte le rappresentazioni delle immagini.

Nei sensori CCD a CMOS, ogni singolo elemento fotosensibile (pixel) è costituito da un elemento che accumula la corrente emessa dal fotodiodo (la superficie attiva del pixel). Ogni unità è connessa singolarmente attraverso un amplificatore allo stadio di uscita.

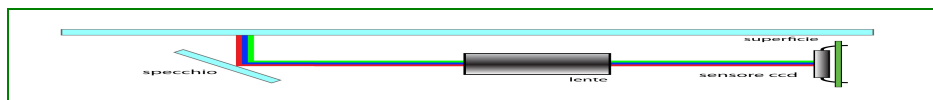



Figura 5.2: schema di un sensore CCD

Sensore CIS (*Contact Image Sensor*). Possiede un'ottica molto semplice, in quanto il sensore è posto subito sotto il vetro e la lampada è sostituita da una serie di led dei tre colori RGB, che si accendono e spengono in rapida sequenza illuminando l'originale. Il CIS registra le variazioni di luminosità, senza usare specchi o lenti; è dunque poco luminoso e ha una scarsa profondità di campo.

Figura 5.3: schema di un sensore CIS

5.3.2 Risoluzione

La risoluzione ottica descrive la quantità di informazioni che il sistema ottico dello strumento di acquisizione può campionare. Indica il numero di pixel per unità di spazio (solitamente il pollice, *inch*). L'unità di misura è nota come *pixel per inch* (PPI) o *dot per inch* (PDI); si misura in pixel per pollice lineare e non pollice quadrato: un immagine a una risoluzione di 300 dpi avrà 300 pixel in orizzontale e 300 pixel in verticale, che corrispondono a 90.000 pixel per pollice quadrato dell'immagine. I fattori che definiscono la risoluzione ottica possono variare a seconda del tipo di periferica. In genere, si considerano il numero dei singoli sensori disposti sulla linea di scansione e la larghezza massima degli originali che lo scanner può accettare. La distanza percorsa dal meccanismo in movimento dello scanner per produrre ogni linea di scansione determina la risoluzione verticale, che può essere talvolta più elevata di quella orizzontale. Le macchine fotografiche digitali, le videocamere e alcuni scanner per trasparenze normalmente utilizzano un dispositivo CCD di forma rettangolare (non un tipo lineare in movimento), che fissa il numero totale di pixel che è possibile catturare in entrambe le dimensioni. Negli scanner a tamburo, la massima risoluzione ottica è determinata dalla velocità di rotazione, la luminosità della sorgente luminosa, le caratteristiche del motore e le dimensioni delle aperture delle lenti.

In genere la risoluzione di uno scanner è definita, quindi, con due misure: *x-dpi* e *y-dpi*, che corrispondono alle coordinate orizzontali e verticali. Esiste una terza misura di risoluzione, detta risoluzione interpolata, una funzione software che permette di ottenere immagini a una risoluzione maggiore di quella che lo scanner in realtà fisicamente consente. Questo non provoca un miglioramento reale della qualità dell'immagine; è bene quindi fidarsi soltanto dei valori di risoluzione effettiva. Per la medesima ragione, una volta che una immagine è stata digitalizzata o creata con una determinata risoluzione, l'incremento manuale della risoluzione non produce un miglioramento della qualità: la stessa informazione viene semplicemente distribuita tra un numero maggiore di pixel.

5.3.3 Profondità di colore

La profondità di colore (*color depht*) indica la quantità di bit necessari per rappresentare il colore di un singolo pixel in un'immagine *raster*. Questa grandezza è rappresentata in *bits per pixel* (BPP). Per determinare con quanti colori può essere visualizzata un'immagine, è sufficiente elevare 2 alla potenza pari del numero di bit associato al tipo di dati.

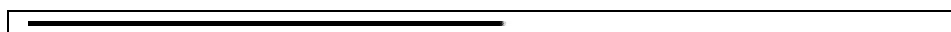



Figura 5.4: un'immagine in bianco e nero (1 bit). Ogni pixel può assumere 2 valori (2^1)

Figura 5.5: un'immagine a 16 colori (4 bit). Ogni pixel può visualizzare 16 colori (2^4).

Figura 5.6: un'immagine in scala di grigi (8 bit). Ogni pixel può assumere 256 differenti valori (2^8).

Nel pixel la tonalità è indicata come combinazione dell'intensità dei tre colori fondamentali RGB. In base al numero di bit assegnato a ciascun canale è possibile stabilire quanti valori diversi può assumere il colore. Con il modello di profondità *truecolor* è possibile riprodurre immagini in modo molto fedele alla realtà, fino a rappresentare più di sedici milioni di colori distinti. Con questa profondità si usano 8 bit per rappresentare il rosso, 8 bit per rappresentare il blu ed 8 bit per rappresentare il verde. I 256 livelli d'intensità per ciascun canale si combinano per produrre un totale di 16.777.216 colori (256 × 256 × 256). Per la maggior parte delle immagini, questa profondità consente di visualizzare sfumature ancora più fini di quelle che l'occhio umano riesce a distinguere. Livelli di profondità tali sono utili nell'acquisizione delle immagini nelle quali la resa di dettaglio deve essere estremamente precisa.

bit	potenza	colori	descrizione
1	2^1	2	Bianco e nero
2	2^2	4	CGA
4	2^4	16	VGA
8	2^8	256	Mezzatinta (toni di grigio) o *Pseudocolor*

16	2^{16}	65.536	*Hicolor*
24	2^{24}	16.777.216	*Truecolor*

Tabella 5.1: Valori tipici della profondità di colore

5.3.4 Gamma di densità

La gamma di densità indica la capacità di riprodurre minime variazioni di tonalità e si esprime con la differenza tra i toni più chiari (*dmin*) e i toni più scuri (*dmax*) che il dispositivo riesce a rilevare. L'intervallo di densità, cioè la differenza dalle densità estreme, descrive il livello minimo e il livello massimo di nero percepibile. Non esistono dispositivi in grado di leggere i valori estremi di nero (0% e 100%). Maggiore è l'intervallo di densità maggiore è il range di tonalità distinguibili.

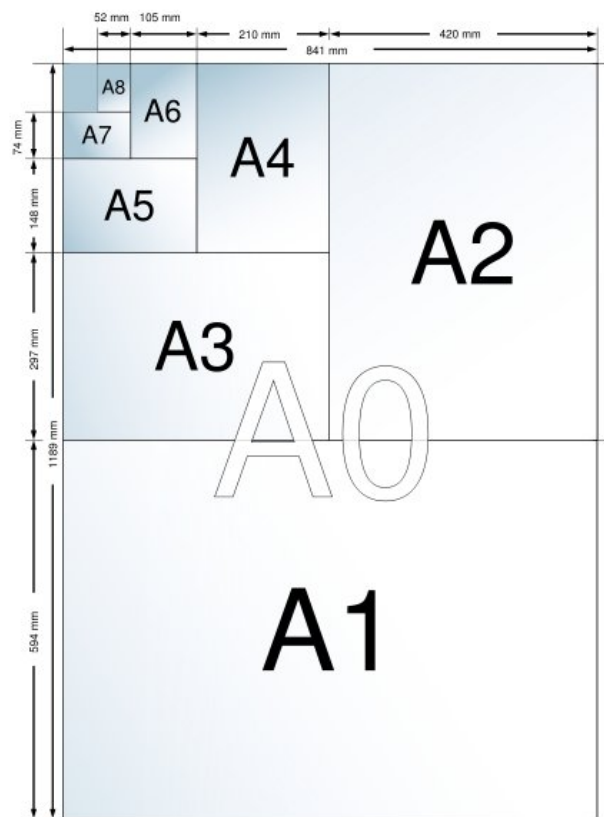



Figura 5.6: proporzioni dei formati di tipo A

Più elevata è la gamma dinamica di uno scanner o di un'originale, maggiore sarà la gamma di livelli di luce che potrà

rilevare, oppure ostruire o assorbire. Un dispositivo per scansioni che possiede un'ampia gamma dinamica è in grado, quindi, di riprodurre maggiori dettagli. Questo aspetto è percepibile soprattutto nelle ombre (le zone più scure delle immagini) dove è più difficile rilevare dettagli e differenziare i vari livelli di luce in quanto esiste una debole energia luminosa che riflette o trasmette i dettagli delle zone scure.

La gamma dinamica può variare anche tra dispositivi che presentano la stessa profondità di bit nominale.

5.3.5 Dimensioni e formato

L'ISO 216[14] è la norma che descrive le dimensioni dei formati standard di carta usati in numerosi paesi, specialmente in Europa. Definisce tre formati fondamentali denominati *A*, *B* e *C*.

Il formato *A* è stato definito partendo dal foglio A0, di superficie pari a 1 m² e di proporzioni tali da ricavare per dimezzamento gli altri formati, mantenendo le stesse proporzioni. Il formato A0 è stato definito in modo tale che le proporzioni iniziali dei lati a e b fossero tali che $\frac{b}{a} = \frac{a}{b/2}$, da cui deriva che $\frac{b}{a} = \sqrt{2}$, che combinata a $ab = 1$ determina le dimensioni del foglio A0: 841x1189 mm (arrotondati).

[14] Questo standard fu inizialmente adottato dal DIN in Germania nel 1922 (è per questo che è ancora oggi comune l'uso di far precedere ai nomi dei singoli formati la sigla DIN) anche se alcuni dei formati furono già sviluppati durante la Rivoluzione Francese per poi essere dimenticati.

Per tutti i formati *A* il rapporto tra i lati del foglio è quindi pari a √2; tale rapporto resta invariato anche se il foglio viene tagliato (o piegato) lungo il suo lato più lungo. Il passaggio da un formato a un altro immediatamente inferiore (per esempio da A4 a A5) si ottiene quindi dimezzando il lato maggiore; viceversa il passaggio verso dimensioni superiori (come da A4 a A3) si ottiene raddoppiando il lato minore.

	Millimetri		**Pollici**		**Punti**	
formato	**lato lungo**	**lato corto**	**lato lungo**	**lato corto**	**lato lungo**	**lato corto**
A1	841	594	33.11	23.39	2384	1684
A2	594	420	23.39	16.54	1684	1190
A3	420	297	16.54	11.69	1190	842
A4	297	210	11.69	8.27	842	595
A5	210	148	8.27	5.83	595	490
A6	148	105	5.83	4.13	420	298

Tabella 5.2: misure dei formati A in millimetri, pollici e punti tipografici.

Il formato dello scanner è un requisito da tenere in grande considerazione. Ogni dispositivo prevede una dimensione massima del foglio da sottoporre a scansione, nei formati standard. È necessario, quindi, valutare con precisione i formati dei documenti che lo scanner potrà essere in grado di gestire, stimando le dimensioni fisiche massime dei documenti da acquisire.

5.3.6 Velocità di scansione

La velocità con cui lo scanner è capace di effettuare la scansione è un parametro importante. Uno scanner precisissimo e di elevatissima qualità, ma con basse prestazioni in termini di velocità, potrebbe essere la scelta meno opportuna per acquisire ingenti quantità di documenti e immagini in tempi ristretti. In ogni caso è conveniente considerare sia l'aspetto qualitivo che la rapidità di scansione per adottare un dispositivo che risponda alle esigenze del progetto di digitalizzazione.

6. Dispositivi di acquisizione ottica

Ogni dispositivo di acquisizione ottica presenta alcune caratteristiche fondamentali. È essenziale, quindi, conoscere le varie tipologie per valutare ponderatamente la scelta del dispositivo, o dei dispositivi, da impiegare per l'acquisizione dei documenti digitali, considerando le prestazioni, i requisiti, ma anche la spesa economica da sostenere.

Di seguito si prenderanno in considerazione le tipologie di scanner più o meno comuni o accessibili, cercando di offrire una panoramica esaustiva.

6.1 Scanner a ripresa in piano

Gli scanner a ripresa in piano (o *flatbed scanner*) utilizzano un gruppo optoelettronico che scorre in direzione parallela al documento, mentre questo rimane immobile su un piano orizzontale. Il gruppo è composto da una lampada fredda che emette luce bianca (nei primi apparecchi era una lampada a neon, oggi sostituita da lampade allo *xenon* o fluorescenti a catodo freddo). Il fascio luminoso avanza proiettando la luce

progressivamente sulla superficie sulla quale è posto il documento ed è raccolto da un sistema di lenti che lo suddivide nelle tre componenti fondamentali RGB. Queste vengono indirizzate ai sensori che possono essere di due tipi: CCD (*Charge Coupled Device*) o CIS (*Contact Image Sensor*). Più sensori possono essere montati sul gruppo optoelettronico. All'avanzare di quest'ultimo, i diversi sensori registrano l'immagine, consentendo di aumentare la risoluzione verticale. Questo valore definisce la risoluzione meccanica (risoluzione-y). L'altra dimensione di risoluzione dipende esclusivamente dalla grandezza del sensore (risoluzione-x).

Figura 6.7 esempi di scanner a ripresa in piano. Modelli Kodak i200 e i260. Fonte: www.scansolutions.co.uk

Uno scanner piano ha costi accessibili, anche nelle sue versioni professionali, ed è abbastanza semplice da usare. Le prestazioni sono versatili e si raggiungono risoluzioni ottiche molto alte. Può essere usato per quella documentazione che non richiede particolari attenzione alla separazione delle pagine o l'apertura a 180 gradi del volume, né particolare sensibilità della carta all'irradiamento da parte di una fonte di luce e calore.

Sensore	CCD – CIS
Risoluzione	300 – 5000 dpi (integrali)
Profondità di colore	24 – 48 bit
Gamma di densità	2,8 - 3,4
Formato	A4 – A3
Costo (euro)	150 - 2500

Tabella 6.3: caratteristiche tecniche di uno scanner piano. Fonte: Space S.p.a

6.2 Scanner planetario (*overhead*)

Lo scanner planetario (o *overhead scanner*) possiede una serie di funzioni fondamentali per produrre in tempi brevi la digitalizzazione di un documento anche molto fragile in maniera fedele. Non presenta, infatti, meccanismi che possano danneggiare gli originali e permette di ottenere scansioni di altissima qualità. È ideale per la digitalizzazione di libri antichi, le cui pagine non possono essere soggette a pressione, o che presentano rilegature in cattive condizioni.

Figura 6.8: esempi di scanner planetari. Modelli CopiBook A2 e DigiBook Suprascan A1. Fonte: http://www.i2s-bookscanner.com/

La ripresa dell'immagine avviene dall'alto, il piano su cui è poggiato il libro è basculante e questo permette di compensare il

dislivello che si crea tra le due pagine affiancate di un volume, in modo da metterle sempre in linea rispetto al piano di ripresa. ll libro è posizionato una volta sola, a faccia in su, e centrato con cura rispetto al piano di ripresa. L'unico movimento a cui è il volume è sottoposto è quello di sfogliare le pagine: non è soggetto a pressioni, poiché semplicemente poggiato sul piano, e un sistema di correzione ottica dell'immagine elimina ombre, curvature e disallineamenti. A fronte di questi vantaggi, bisogna considerare gli altissimi costi di questo prodotto.

Sensore	CCD
Risoluzione	300 – 1600 dpi (integrali)
Profondità di colore	24 – 48 bit
Gamma di densità	3,2 – 4
Formato	A4 – A0
Costo (euro)	25000 – 40000 (scanner 600 ppi)

Tabella 6.4 caratteristiche tecniche di uno scanner planetario. Fonte: Space S.p.a

6.3 Fotocamera con stativo

La qualità di acquisizione di una fotocamera digitale può raggiungere livelli davvero elevati, sia in termini di risoluzione che di fedeltà dei colori. Insieme a uno stativo, un sostegno verticale dotato di tre gambe richiudibili a ombrello, questa soluzione può rappresentare un'ottima alternativa allo scanner piano e al planetario, per opere ordinarie. La macchina fotografica, preferibilmente una reflex digitale con ottica adeguata, è sistemata

sul supporto in posizione perpendicolare al documento originale, con un sistema di illuminazione formato da due o più lampade laterali. È possibile, inoltre, utilizzare un telecomando remoto esterno per comandare gli scatti, in modo tale da evitare ogni tipo di spostamento accidentale manovrando la fotocamera.

Figura 6.9: esempi di stativi IFF. Modelli IFF Mini Repro 2 FF1192 e Tresor CS-1070. A destra, una fotocamera reflex digitale Canon Eos 450d. Fonte: http://www.tuttofoto.com

La qualità globale è garantita e i costi sono contenuti rispetto all'acquisto di uno scanner planetario. La qualità del file digitalizzato è stabilita dalla risoluzione della fotocamera e può essere sufficiente per progetti che non abbiano necessità di elevate risoluzioni di scansione (fino a 300 dpi).

Oltre al numero di pixel registrati dalla fotocamera, è fondamentale considerare anche le dimensioni del sensore: all'aumentare di questo, migliora la qualità del segnale registrato e il rapporto segnale/rumore. Questo fenomeno si evidenzia in modo particolare nelle riprese a bassa luminosità dove possono comparire

delle imprecisioni di immagine, dovute a segnali derivanti dal rumore elettrico di fondo degli elementi fotosensibili.

Fotocamera	Risoluz. in megapixel	Dimensione immagine
Olympus E-400	10.8	3648 x 2736
Canon eos 400d	10,1	3888 x 2592
Nikon D40x	12,1	3872 x 2592

Tabella 6.5: comparazione di tre fotocamere reflex.

Il risultato di un processo di digitalizzazione attraverso una fotocamera con il supporto di uno stativo è positivo per la qualità di immagine e la velocità di esecuzione, ma insufficiente per la correzione delle imperfezioni ottiche di ripresa, come la curvatura della pagina e la comparsa di ombre. Ogni immagine prodotta deve essere corretta manualmente intervenendo su centratura, contrasto, pulizia dei bordi e altri fattori.

Sensore	CCD – CMOS
Risoluzione	10 megapixel in su
Profondità di colore	24 – 48 bit
Gamma di densità	2,8 - 3,4
Formato	A4 – A3
Costo (euro)	600-1500

Tabella 6.6: caratteristiche tecniche di una fotocamera digitale da utilizzare per la digitalizzazione

6.4 Scanner a tamburo

Negli scanner a tamburo (o *drum scanner*) il gruppo optoelettronico è fermo rispetto al documento, mentre questo è

disteso intorno a un rullo rotante. Un fascio di luce bianca è proiettato sul documento e indirizzato al sensore, che separa la luce riflessa dalla pagina nelle tre componenti fondamentali RGB. Ciascuna componente è poi inviata a un fotomoltiplicatore (PMT, *Photo Multiplier Tube*) che trasforma il raggio luminoso in un segnale elettrico.

Figura 6.10: esempi di scanner a tamburo. Modelli Fuji Celsis 6250 e AZTEK Premier Drum Scanner. Fonte: www.aztek.com

Gli scanner a tamburo sono caratterizzati da un alto livello qualitativo e grande produttività, ma hanno costi di acquisto proibitivi. Sono consigliati solo per esigenze di grande precisione e acquisizione ad altissima risoluzione e non permettono la scansione di volumi o documenti rilegati. Il processo di rotazione del rullo è pericoloso, inoltre, per fogli fragili o in avanzato stato di deterioramento.

Sensore	PMT
Risoluzione	12000 dpi
Profondità di colore	24 - 48 bit
Gamma di densità	3,6 - 4 (dmax 4-4,2)
Costi (euro)	100.000 +

Tabella 6.7: caratteristiche di uno scanner a tamburo. Fonte: Space S.p.A.

6.5 Scanner a inserimento di fogli

Gli scanner a inserimento di fogli (o *sheetfeed scanner*) funzionano per mezzo di un gruppo optoelettronico fisso e un sistema di rulli che fa scorrere ciascun foglio davanti al gruppo stesso. Durante questo movimento, i documenti sono investiti da un raggio luminoso; la luce riflessa è indirizzata tramite le lenti sui sensori e trasformata in segnale elettrico. Questo tipo di scanner è dotato di un vassoio entro il quale può essere inserito un numero elevato di fogli e spesso presenta due sensori installati, che consentono di acquisire entrambi i lati di un foglio in un singolo passaggio.

Figura 6.11: esempi di scanner a inserimento di fogli. Modelli Xerox DocuMate 262 e HP N6010 Scanjet. Fonte: http://www.itreviews.co.uk/

Questo sistema può raggiungere velocità di scansione molto elevate ed è utile per digitalizzare grandi quantità di documenti in tempi ridotti. Di contro, può acquisire soltanto documenti in generale non superiori all'A3; le meccaniche, inoltre, possono danneggiare irrimediabilmente fogli delicati. Rimane una soluzione

valida per quei documenti dai quali si voglia estrarre testo attraverso un sistema automatico di ricognizione dei caratteri (OCR).

Sensore	CCD - CIS
Risoluzione	1200 dpi
Profondità di colore	24 - 48 bit
Gamma di densità	2,8 – 3,4
Formato	A5 – A3
Costi (euro)	250 - 1000

Tabella 6.8: caratteristiche di uno scanner a inserimento fogli.

6.6 Scanner per diapositive e pellicole

Gli scanner per diapositive e pellicole sono finalizzati ad acquisire digitalmente il contenuto di pellicole fotografiche (negativi o diapositive). In questo caso il gruppo optoelettronico e i sensori non percepiscono la luce riflessa, ma quella che viene proiettata attraverso la pellicola stessa.

Figura 6.12: esempi di scanner per pellicole e dispositive. Modelli VuPoint Solutions FS-C1-VP e Plustek OpticFilm 7200 Fonte:http://www.pixmania.com

Sono caratterizzati da un alta risoluzione ottica, che consente di effettuare un sufficiente ingrandimento di piccoli originali e

dettagli. Le diapositive e le pellicole possiedono un intervallo di densità compreso tra 2,8 e 3,0, quindi, per acquisire tutte le tonalità esistenti, tali scanner devono adottare sensori CCD di alta qualità.

Sensore	CCD
Supporti accettati	diapositive 35mm, pellicole 6x6, 6x9, 20x12 negative o positive
Risoluzione	1800 - 8000 ppi
profondità di colore	24 - 48 bit
Gamma di densità	2,5-4
Costi (euro)	1000 +

Tabella 6.9: caratteristiche di uno scanner per diapositive e pellicole. Fonte: Space S.p.A.

6.7 Dorsi digitali e a scansione

I dorsi digitali sono l'equivalente di fotocamere digitali, ma molto più potenti: contenengono sensori CCD-CMOS di grosse dimensioni e permettono una acquisizione ad altissima risoluzione. Si utilizzano tramite adattatori su macchine fotografiche di medio e grande formato. Esistono tre tipologie di dorsi digitali: a scansione, *one-shot* e *multi-shot*. Nei dorsi *multi-shot*, la fotografia è composta attraverso l'unione di tre esposizioni differenti, una per ogni colore della terna RGB. Al contrario, nei dorsi *one-shot* l'immagine è prodotta in un'unica esposizione.

Figura 6.13: esempi di dorsi digitali. Modelli Leaf AFi 7 e Hasselblad Ixpress V96C. Fonte: <http://www.sabatini.ws>

I dorsi a scansione sono usati in genere con fotocamere di grande formato, e funzionano in modo analogo a uno scanner piano. Una singola riga di fotoricettori è mossa da un motorino lungo il dorso, producendo come risultato un'immagine finale di elevatissima risoluzione. Il tempo di scatto si moltiplica per la velocità impiegata dal sensore a coprire l'intero campo di acquisizione; per questo motivo i dorsi a scansione possono essere utilizzati solo in studi appropriati e in situazioni di luce controllata.

Sensore	CCD
Risoluzione	22+ megapixel
Spazio colore	RGB, ICC, CMYK
Dimensione pixel	9 X 9 micron
Costi (euro)	20.000+

Tabella 6.10: caratteristiche di un dorso digitale. Fonte: vari.

6.8 Sistemi di scansione per tipologia di documento

Come già illustrato nel capitolo 4, esistono differenti tipi di materiali che possono essere sottoposti a scansione. Ciascuno di

essi richiede un'appropriata strategia di digitalizzazione al fine di ottenere risultati apprezzabili.

Di conseguenza, oltre alle potenzialità tecniche, la scelta del dispositivo di scansione deve essere effettuata in base alle caratteristiche degli originali da acquisire: tipologie, condizioni fische e tipi di supporto.

Libri semplici. La scansione di questo tipo di materiale non richiede l'utilizzo di strumentazioni sofisticate o costose. Infatti, ciò che fondamentalmente interessa è l'estrazione del testo, non l'acquisizione delle immagini o altre qualità. Uno scanner piano può essere più che sufficiente in termini qualitativi; a livello di rapidità di lettura, nel caso di documenti non rilegati, uno scanner a inserimento di fogli può ridurre significamene i tempi di acquisizione e automatizzare i processi.

Libri manoscritti, libri di architettura, pergamene. Data la fragilità dei documenti e l'importanza dell'immagine, sarà necessario un sistema ad alte prestazioni. La soluzione migliore è lo scanner planetario: altri tipi di dispositivi non sono in grado di trattare adeguatamente i supporti delicati (come lo scanner a tamburo), oppure operano esercitando pressioni fisiche sul documento o ancora sono insufficienti in termini qualitativi (come lo scanner piano). Una soluzione più economica al planetario è

rappresentata dalla fotocamera digitale con il supporto di uno stativo. Per i manoscritti e le pergamene possono essere impiegati sistemi di protezione specifici, come teche di vetro dette *sandwich* che permettono di srotolare i documenti avvolti su se stessi senza subire pressioni dannose.

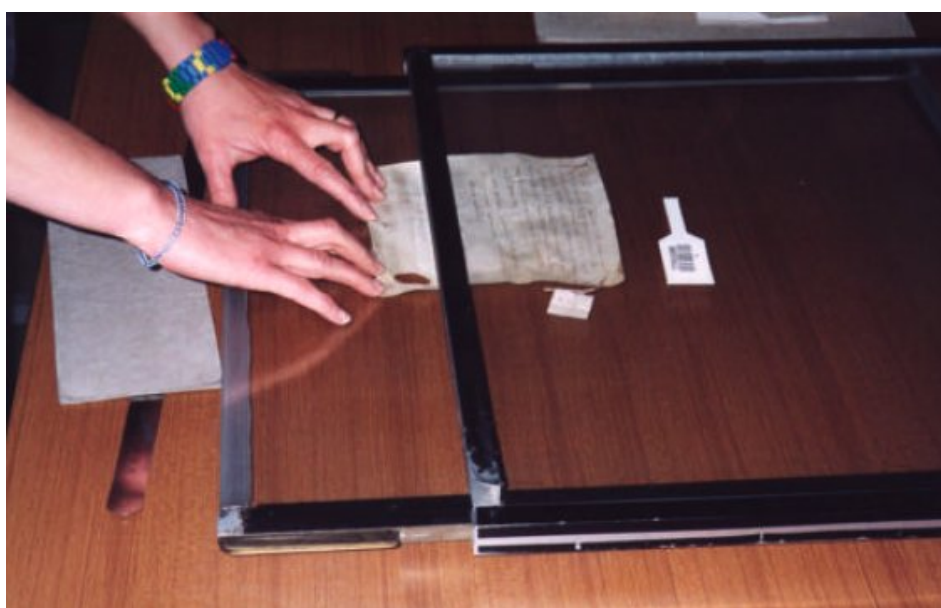



Figura 6.14: telaio scorrevole in vetro a "sandwich" per il posizionamento e la ripresa di documenti arrotolati

Bandi, manifesti e fogli volanti. Considerata la natura eterogenea della tipologia di documenti, composta da singoli fogli di dimensioni variabili, è possibile utilizzare sia scanner piani che scanner planetari. Se le opere presentano in alcuni casi dimensioni maggiori rispetto al tipo di strumento adoperato, e non si ha la possibilità di utilizzare strumenti alternativi, si può procedere a

digitalizzare porzioni di documento per poi unire le singole parti. Questa tecnica è detta *a mosaico*.

Fotografie. Per questa tipologia di documenti esiste un'ampia gamma di soluzioni: la scelta migliore per una acquisizione di alta qualità e risoluzione è la digitalizzazione per mezzo di scanner a tamburo, ma per velocizzare le attività possono essere impiegati anche scanner piani e planetari, seppur con risultati inferiori.

Opere iconografiche. È necessaria una strumentazione di alto profilo, poiché è necessario acquisire molti dettagli e di conseguenza una mole di dati considerevole. La scansione digitale deve spesso avvenire non soltanto nello spettro del visibile, ma anche nell'ultravioletto e nell'infrarosso. Si utilizzano allo scopo dorsi digitali con filtri appositi applicati alle fotocamere; le opere devono essere illuminate in maniera adeguata, in modo da registrare informazioni sulla radiazione emessa.

7. Software

Il funzionamento di un dispositivo di acquisizione ottica non dipende soltanto dalla qualità hardware dello strumento. È determinante utilizzare programmi informatici appropriati e di qualità per ottenere un risultato di buona qualità nell'acquisizione delle immagini o per operazioni successive.

7.1.1 Driver

Il *driver* dello scanner, il programma che "pilota" e funge da interfaccia tra il computer e il dispositivo, permette al sistema operativo di utilizzare la periferica dialogando attraverso un'interfaccia standard, i registri del controllore della periferica. Senza addentrarsi troppo nelle questioni informatiche, che variano da caso a caso, è bene accertarsi che lo scanner sia corredato da *driver* funzionanti e compatibili con il sistema operativo da utilizzare.

7.1.2 Twain

Per permettere allo scanner di rapportarsi perfettamente con tutti i programmi di acquisizione dell'immagine, è stata messa a

punto un'API (*Application Programming Interface*, un software che fa da interfaccia tra altri software diversi per migliorarne l'interazione), denominata *twain*[15].

Il *twain* è uno standard di comunicazione e permette di pilotare lo scanner e l'acquisizione all'interno di altri software di grafica o di impaginazione che supportano questo protocollo. In questo modo si accelera l'operatività e si migliora la qualità e l'efficienza delle prestazioni dello scanner, perché immediatamente si può avere un'idea del risultato del lavoro di acquisizione già all'interno del contesto in cui il file sarà inserito.

[15] La parola *twain* deriva da una frase dal racconto di Kipling *The Ballad of East and West*: «and never the twain shall meet» con la quale si volle sottolineare la difficoltà a quei tempi di collegare gli scanner ai personal computer. Si decise di scriverla in maiuscolo per renderla più riconoscibile. Questo particolare la fece apparire come un acronimo e molti si dilettarono a trovarne uno efficace. Infine non se ne adottò alcuno ufficialmente ma quello che va per la maggiore è *Technology (o Toolkit) Without An (o Any) Interesting (o Important) Name*, letteralmente "tecnologia senza un nome importante".

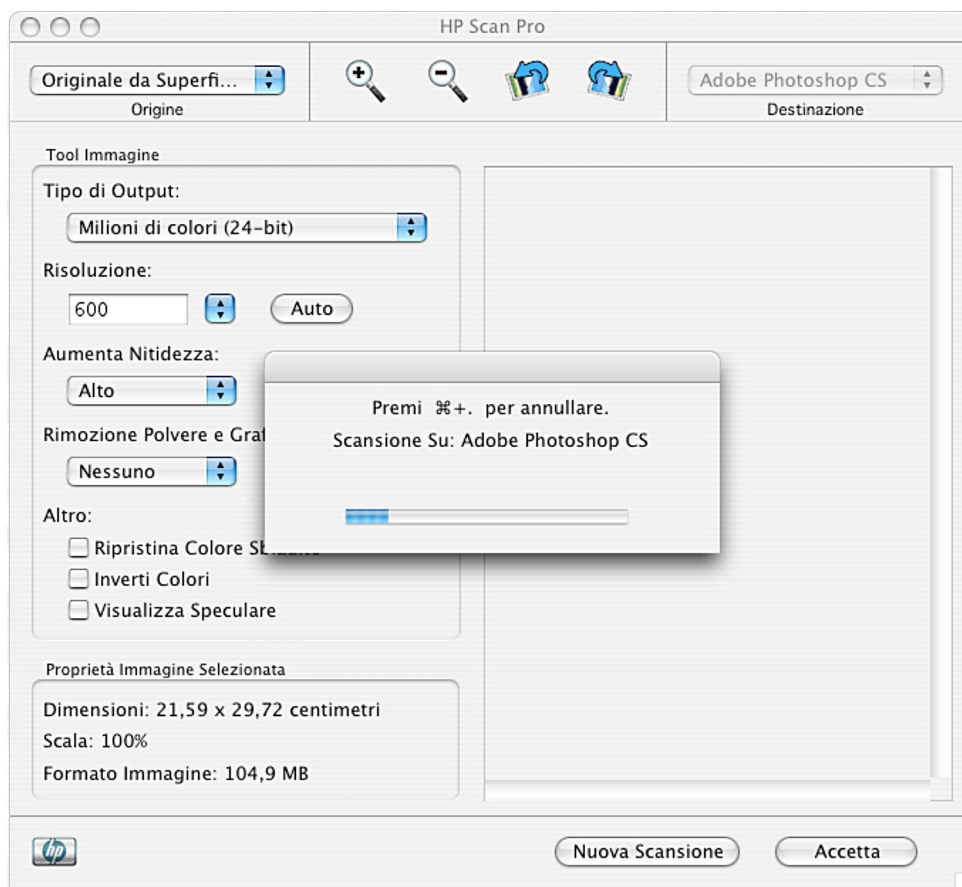



Figura 7.15: interfaccia *twain* di scansione tra Adobe Photoshop CS e Scanner Hp

7.1.3 Fotoritocco

Se i software di scansione sono già in grado di produrre un'immagine accettabile, la maggior parte degli interventi più mirati sono eseguiti attraverso alcuni software per il fotoritocco. Questa operazione si svolge nella fase successiva all'acquisizione, in post produzione. Sarà affrontata dettagliatamente in seguito nel capitolo 11, dedicato all'*editing* delle immagini acquisite.

Gli scanner sono in genere corredati da un CD ROM in dotazione che comprende vari tipi di software adatti al ritocco di immagini digitali. Naturalmente si tratta di programmi piuttosto semplici o versioni limitate di programmi maggiori. La scelta migliore è quella di affidarsi a un software più completo e sviluppato, valutandone ovviamente la spesa. Ad ogni modo, esistono programmi molto avanzati e completamente gratuiti scaricabili da internet.

7.1.4 OCR (*Optical Character Recognition*)

In genere, lo scanner acquisisce un documento di testo in un formato digitale identico a quello di una qualsiasi altra immagine. Il risultato è quindi una "fotografia" del testo e i caratteri (lettere e numeri) contenuti non sono altro che un insieme di pixel a una certa risoluzione e con una certa profondità di colore. In queste condizioni è quindi possibile osservare il testo, ma non apportarvi modifiche. È spesso necessario riuscire a trasformare questi files in testo editabile, dello stesso tipo di quelli prodotti da un software per la videoscrittura. Attraverso i sistemi di *Optical Character Recognition* (riconoscimento ottico dei caratteri) è possibile convertire un'immagine contenente testo, generalmente acquisita tramite scanner o altri dispositivi ottici complessi, in testo digitale modificabile con un normale editor. Il testo può essere convertito in formato ASCII semplice, Unicode o, nel caso dei sistemi più

avanzati, in un formato contenente anche gli elementi tipografici di formattazione del documento.

I software OCR sono spesso inclusi nel pacchetto software in dotazione con il dispositivo, ma possono anche essere acquistati a parte. Quelli più evoluti riescono a riconoscere quasi il 100% dei caratteri di un testo, a condizione che la pagina originale non sia ovviamente troppo deteriorata. Ad ogni modo, utilizzando un buon programma di fotoritocco si può migliorare fortemente la qualità finale del documento acquisito e far sì che questo sia più facilmente riconoscibile dall'OCR.

7.1.5 Database

La gestione di centinaia o migliaia di immagini acquisite in formato digitale può richiedere l'utilizzo di un database, strutturato in modo tale da consentire la gestione dei dati stessi (l'inserimento, la ricerca ecc.).

La banca dati è capace di contenere molte informazioni e di essere gestita direttamente dai software utilizzati, interfacciandosi con il sistema operativo.

8. Formati di output

8.1 La compressione

L'obiettivo della compressione delle immagini è quello di ridurre quanto più possibile il numero di bit necessari per la rappresentazione e la ricostruzione di un duplicato identico (o quanto meno fedele) dell'immagine originale. Al fine di ottimizzare gli spazi o la divulgazione in rete, le immagini possono essere compresse, ossia ridotte nelle dimensioni finali rispetto un formato "grezzo", corrispondente alla memorizzazione diretta dei tre valori RGB per ogni pixel.

Le tipologie di compressione possono essere divise in due grandi classi: senza perdita di qualità e con perdita di qualità.

La tecnica di compressione che consente di recuperare esattamente i dati originali viene detta *lossless* (senza perdita di informazione), in contrapposizione alle tecniche *lossy* (con perdita di informazione), che permettono di riottenere solo copie approssimativamente identiche all'originale. Entrambe le tecniche si basano sul fatto che nelle immagini i campioni adiacenti hanno valori di intensità simili ed è possibile trovare significative e importanti correlazioni spaziali; se tali correlazioni sono trattate in

modo appropriato, il numero totale di bit può essere drasticamente ridotto: l'immagine è partizionata in una serie di regioni, ognuna dotata di caratteristiche peculiari, separate tra di loro da contorni, in modo che tali regioni coincidano il più fedelmente possibile con gli oggetti presenti nella scena di interesse. La maggiore efficienza di compressione che ne può conseguire deriva sia dall'esistenza di metodi molto efficienti per la codifica dei contorni, sia dal fatto che le regioni all'interno dei contorni hanno in generale un contenuto informativo molto basso e possono quindi essere rappresentate da un numero relativamente basso di bit. L'inconveniente principale di tali tecniche risiede tuttavia nella loro enorme complessità computazionale, che risulta per di più legata in maniera essenziale al contenuto informativo dell'immagine.

Gli algoritmi di compressione *lossless* conservano le informazioni originali del file compresso; tramite tali informazioni è possibile risalire successivamente alla struttura originale del file in ogni sua parte.

Gli algoritmi *lossy* più comunemente usati sono quelli di *transform coding*. Tale metodo consiste essenzialmente nel trasformare l'insieme di pixel in un altro insieme di coefficienti che presenti una correlazione minore, e che possa dunque essere codificato in maniera più efficiente, sebbene comporti una perdita di dati a fronte di un minor peso in termini di kilobytes.

8.2 I formati

I dati di tipo *raster* possono essere memorizzati attraverso tipologie di file che, sfruttando algoritmi di compressione diversi, gravano in modo differente sul supporto di memorizzazione: più si comprime un'immagine, più il file si riduce di dimensioni in kilobytes e più si riduce la qualità dell'immagine stessa.

8.2.1 RAW

Il formato RAW (*Read After Write*) è un particolare metodo di memorizzazione dei dati descrittori di un'immagine, che permette di non avere perdite di qualità della registrazione su un qualsiasi supporto rispetto ai segnali catturati dal sensore e successivamente composti per interpolazione dal processore d'immagine della fotocamera nelle sue tre componenti fondamentali RGB. La risoluzione massima reale dell'immagine rimane quella determinata dalle caratteristiche del sensore. Il formato RAW non possiede una vera e propria estensione, ovvero non esiste un file ".raw", poiché ogni produttore ha un formato RAW proprietario.

Prive di fattori di compressione (in inglese *raw* significa "grezzo"), le informazioni rappresentano esattamente ciò che è prodotto sul sensore digitale.

profondità colore	compressione	supporto multi-immagine
fino a 48 bbp	nessuna	no

Tabella 8.11: tabella riassuntiva delle caratteristiche del formato RAW

8.2.2 BMP

Il BMP (*Microsoft Windows Bitmap*) è il formato di visualizzazione delle immagini dei sistema operativo Windows e, sebbene sia uno dei formati più vecchi, è ancora molto utilizzato e riconosciuto da qualunque piattaforma. Esiste, anche se non molto usata, la possibilità di comprimere le bitmap da 16 e 256 colori con l'algoritmo RLE. La compressione RLE non è tuttavia efficiente come altri metodi di compressione *lossless*. Le immagini bitmap, anche quelle compresse, occupano di regola più spazio su disco rispetto ad altri formati raster e sono quindi poco meno adatte alla trasmissione di immagini via Internet o alla memorizzazione permanente.

profondità colore	compressione	supporto multi-immagine
fino a 24 bbp	RLE, nessuna	no

Tabella 8.12: tabella riassuntiva delle caratteristiche del formato BMP

8.2.3 GIF

Il formato GIF (*Graphics Interchange Format*) è stato introdotto negli anni Ottanta come metodo efficiente di trasmissione delle

immagini su reti di dati. GIF usa una forma di compressione LZW, che consiste nel ridurre il numero di colori al numero minimo necessario e nell'eliminare i colori isolati non necessari per la rappresentazione dell'immagine. Lo svantaggio risiede nella profondità colore a 8 bit, che consente di usare una tavolozza di 256 colori, decisamente limitata per una minima apprezzabilità di colore in una immagine che abbia pretese di fedeltà all'originale.

profondità colore	compressione	supporto multi-immagine
fino a 8 bbp	LZW	sì

Tabella 8.13: tabella riassuntiva delle caratteristiche del formato GIF

8.2.4 TIFF

Il formato digitale TIFF (*Tagged Image File Format*) permette di rappresentare immagini con diversi spazi di colore: scale di grigio, RGB, CMYK, CIELab. Un file TIFF può essere aperto, elaborato e salvato tutte le volte che si desidera, senza che intervengano cadute di qualità dell'immagine. Offre inoltre il vantaggio di essere riconosciuto da tutti i più diffusi programmi di elaborazione delle immagini.

Lo svantaggio dell'utilizzo del formato TIFF è la pesantezza in termini di megabyte, che influisce anche sui tempi di registrazione delle immagini e di manipolazione delle stesse al computer. Le immagini salvate in TIFF richiedono, quindi, molto spazio per essere memorizzate e non si prestano all'utilizzo sul web.

profondità colore	compressione	supporto multi-immagine
fino a 32 bbp	Nessuna, RLE, LZW, CCITT, JPG	sì

Tabella 8.14: tabella riassuntiva delle caratteristiche del formato TIFF

8.2.5 JPEG

Il formato JPEG (*Joint Photographic Experts Group*) usa una compressione di tipo *lossy*: per diminuire la dimensione dell'immagine alcuni pixel sono eliminati, per poi essere ricostruiti (non esattamente uguali) in fase di decompressione; l'immagine finale non è più, quindi, quella originale. Il formato JPEG è stato tuttavia progettato in modo tale che i pixel eliminati siano quelli meno percettibili all'occhio umano: questo è meno sensibile alle minime variazioni di colore che alle piccole variazioni di luminosità; così se JPEG recupera due pixel di minima variazione di luminosità li conserva, ma se trova due pixel di minima variazione di colore ne conserva in memoria uno soltanto.

profondità colore	compressione	supporto multi-immagine
fino a 24 bbp	DCT	no

Tabella 8.15: tabella riassuntiva delle caratteristiche del formato JPG

8.2.6 Jpeg2000

La tecnica che utilizza JPEG 2000 è analoga alla compressione JPEG. Tuttavia, mentre JPEG utilizza l’algoritmo DCT (*Discrete Cosine Transform*), JPEG2000 utilizza il DWT (*Discrete*

Wavelet Transform). Il formato essere impostato sia come *lossy* che come *lossless*; in caso di compressione, degrada la qualità dell'immagine molto meno del classico JPEG. Può generare immagini compresse scalabili sia in risoluzione sia in livello di dettaglio, lasciando la libertà di scegliere quanta informazione e quali parti dell'immagine utilizzare per la decompressione.

Una delle caratteristiche più importanti dello standard JPEG2000 è la possibilità di definire aree dell'immagine che saranno codificate con qualità più alta rispetto al resto dell'immagine. Queste aree sono dette *regions of interest* (ROI), mentre le parti rimanenti sono indicate come *background* (sfondo).

profondità colore	compressione	supporto multi-immagine
fino a 24 bbp	Nessuna, DWT	no

Tabella 8.16: tabella riassuntiva delle caratteristiche del formato JPG2000

8.2.7 Djvu

DjVu è una tecnologia di compressione d'immagine, un formato file e una piattaforma per la distribuzione via rete di documenti. Si fonda sul modello MRC, uno standard di formato aperto (grazie al progetto DjVuLibre), che propone la segmentazione dell'immagine in livelli, ciascuno dei quali viene compresso con algoritmi specifici. Separando elementi di sfondo, texture della carta e foto e comprimendoli con un algoritmo di tipo

wavelet (lo stesso alla base del formato JPEG 2000), Djvu può garantire un'elevata risoluzione del livello del testo. Con questa tecnica si possono ottenere file molto leggeri, veloci da visualizzare anche su computer di vecchia generazione.

All'estero già moltissimi istituiti di ricerca, biblioteche, archivi, agenzie governative e società hanno deciso di rendere accessibili in rete i loro fondi documentali adottando questo formato; in Italia è ancora poco diffuso.

profondità colore	compressione	supporto multi-immagine
fino a 24 bbp	IW44, JB2	sì

Tabella 8.17: tabella riassuntiva delle caratteristiche del formato djvu

8.2.8 Il formato di descrizione pagina PDF

Il PDF (Portable Document Format) è uno dei più importanti formati per la creazione e la visualizzazione di documenti elettronici. Introdotto dalla casa di produzione Adobe all'inizio degli anni Novanta, ha avuto subito una buona diffusione grazie alla sua grande versatilità.

L'elemento chiave del PDF è il supporto multipagina e la possibilità di inclusione di diversi tipi di files. Permette di conservare, indipendentemente dal creatore o dal destinatario, tutti i font, gli elementi grafici e l'impaginazione del documento originale,

senza che esso venga alterato; è altrettanto efficace nella raccolta e nell'ordinamento di immagini raster.

I file in questo formato, compatti e leggeri, possono essere facilmente visualizzati da qualunque computer, grazie alla possibilità di scaricare gratuitamente il software di lettura Adobe Acrobat Reader dal sito della casa di produzione.

Un file PDF può essere modificabile o meno, in base alla scelta di protezione imposta al documento. Esiste la possibilità di proteggerlo in modo da renderlo, ad esempio, non stampabile o stampabile solo in parte. Si può anche proteggere il testo da modifiche o si può proibire di estrapolarne alcune sue parti.

9. Definizione dei livelli qualitativi

Nello stabilire le impostazioni di scansione per i documenti da digitalizzare, è bene prevedere l'esportazione finale di diverse copie destinate ai vari usi. I documenti master, combinati insieme a quelli consultabili in sede e a quelli visionabili online, devono possedere caratteristiche specifiche e diverse, tali da garantire una fruizione adeguata in base alle circostanze e ai diversi possibili impieghi.

In generale, devono essere custodite copie degli stessi documenti a qualità diverse. L'accesso online presuppone la consultazione di files a risoluzione media, non troppo pesanti in termini di kilobytes, per consentire un trasferimento rapido. Diversamente, la consultazione in sede d'archivio o di biblioteca deve permettere l'osservazione di documenti qualitativamente migliori per avere la possibilità di indagare sugli elementi distintivi anche più minuti. Infine, è sempre necessario conservare delle copie di sicurezza dei files master, a qualità nativa e senza compressione, per garantire l'interoperabilità con possibili piattaforme e standard futuri.

Esattamente dieci anni fa, sono stati formalizzati degli standard qualitativi per le diverse copie di conservazione, ai quali i progetti di digitalizzazione dovrebbero attenersi. Sebbene si tratti di disposizioni vecchie di un decennio, considerati i rapidi sviluppi delle tecnologie informatiche, costituiscono delle linee generali sempre valide che, con qualche piccolo accorgimento, possono essere regolarmente accettate per determinare gli standard dei livelli qualitativi di ogni progetto di digitalizzazione.

9.1 La normativa ICCD del 1998

La *Normativa sull'acquisizione digitale* dell'Istituto Centrale per il Catalogo e la Documentazione (ICCD) del Ministero per i Beni e le Attività Culturali distingue tre livelli qualitativi di immagini digitali, denominati rispettivamente *A*, *B* e *C*.

Il primo livello è quello che riguarda i files master, mentre gli altri due riguardano i derivati di qualità inferiore. Il discorso più interessante è proprio sul master digitale. L'ICCD descrive il livello qualitativo A come

> Immagini ad altissima risoluzione spaziale, da utilizzare essenzialmente per la stampa e come riferimento digitale di alta

qualità dell'originale fotografico (ottenibile con l'uso di scanner professionali)[16].

Il primo livello prevede come formato si output di riferimento il formato PNG, uno standard aperto e *lossless*. Le considerazioni dell'ICCD sono del 1998 e non sono state più aggiornate. Infatti, attualmente lo standard *de facto* per le immagini master generate da scanner è il TIFF che, nonostante sia un formato proprietario della casa di produzone Adobe, è liberamente usato in tutte le scansioni. È consigliabile, quindi, attenersi non tanto a disposizioni di dieci anni fa, ma alla tendenza attuale.

Per il livello B, definito dal Ministero come

> Immagini di media risoluzione spaziale, destinate essenzialmente alla normale consultazione e a corredo di tutte le tipologie di schede (ottenibile con l'uso di strumenti di categoria commerciale)[17].

è previsto l'utilizzo di JPEG e PCD (formato proprietario Kodak poco condiviso), anche se attualmente lo stato della tecnica si dirige verso il JPEG 2000, più recente e dalle migliori prestazioni.

Infine per il livello C, che identifica

---

[16] Ministero per i Beni e le Attività Culturali, Istituto Centrale per il Catalogo e la Documentazione, *Normativa per l'acquisizione digitale delle immagini fotografiche*, 1998<http://www.iccd.beniculturali.it/Catalogazione/standard-catalografici/aquisizione-digitale-delle-immagini-fotografiche/>, pag. 7.

[17] *Ibid.*, pag. 9.

Immagini "francobollo", da utilizzare per la rappresentazione schematica su schermo […] ricavato riducendo via software le immagini dei livelli precedenti[18].

è ancora una volta previsto l'utilizzo dello standard JPEG come formato di output che, per il caso specifico, è ancora attualmente valido e consigliabile.

	Dimensioni	Profondità colore		Formato di output	
		colore	scala di grigi	primario	alternativi
A	3072≅ 3072			png	pcd (kodak)
B	1024≅ 1024	24 bbp	8 bbp	jpg	pcd (kodak), png
C	120≅ 120			jpg	-

Tabella 9.18: tabella riassuntiva degli standard previsti della *Normativa per l'acquisizione digitale delle immagini fotografiche del 2008*. In rosso, le caratteristiche superate dagli standard attuali o ormai desuete.

9.2 Un aggiornamento sui livelli qualitativi

Per ridefinire i livelli qualitativi previsti dieci anni fa dalla normativa, è comunque necessario tenere in considerazione i criteri da essa definiti:

- il formato utilizzato deve essere di larga diffusione;
- deve permettere la rappresentazione di immagini sia a colori reali, sia a scala di grigi;

[18] *Ivi*, pag. 11.

- l'estensione applicata al nome del file deve permettere una rapida e sicura individuazione del formato utilizzato;
- il formato utilizzato (almeno quello principale) deve essere di utilizzo libero (non "proprietario"), in altre parole non deve essere gravato da oneri derivanti da diritti d'autore[19].

Nel rispetto delle le linee guida fornite dal Ministero nel '98 e gli standard *de facto* attualmente condivisi, si può quindi prevedere una quadripartizione per i livelli qualitativi da destinare ai vari usi:

1. una copia ad **altissima risoluzione**, in formato TIFF non compresso, con una risoluzione di 600 dpi rispetto all'originale e con una profondità di colore di 24 bit, da conservare come **copia di sicurezza** in archivio;

2. una copia ad **alta risoluzione** in formato JPEG2000, con una risoluzione di 300 dpi rispetto all'originale e con una profondità di colore di 24 bit per la **consultazione in sede** del formato digitale;

3. una copia a **bassa risoluzione,** in formato JPEG, con una risoluzione di 72 dpi rispetto all'originale e con una profondità di colore di 24 bit per la **consultazione online** del documento;

[19] *Ivi*, pag. 5-6.

4. una **miniatura** in formato JPEG, con una risoluzione di 72 dpi, formato ridotto di massimo 150 pixel per lato e con una profondità di colore di 16 bit per la associazione ai **record catalografici.**

Lo schema appena presentato si propone solo come un suggerimento ed è quindi solo indicativo. Nulla vieta che, per l'utilizzo finale delle immagini, possa essere utilizzato un formato diverso, se questo è più rispondente a particolari esigenze. Tuttavia, questa quadripartizione ha permesso di ottenere risultati apprezzabili durante il progetto di digitalizzazione del fondo manoscritto De Roberto, a quello della digitalizzazione del *Vojage Pittoresque de Sicilie, Malta et Lipari di Houel,* alla digitalizzazione dei materiali d'archivio del documentarista Ugo Saitta che saranno analizzati nella seconda parte di questo lavoro.

10. La pratica della scansione

10.1 Preparazione degli strumenti: la calibrazione

Calibrare un dispositivo informatico equivale a portarlo da una situazione "nativa" di partenza, costituita dalle impostazioni prestabilite dalla casa di produzione del prodotto, a una situazione finale precisamente stabilita, cioè a uno stato noto definito dall'utente. La calibrazione è un processo che precede la digitalizzazione e che si applica sui dispositivi di acquisizione ottica e di visualizzazione a schermo. Ogni dispositivo proveniente da diversi produttori, o tra differenti sistemi dello stesso produttore, possiede un proprio modo di acquisire l'informazione e interpretare le tonalità dei colori digitalizzati. L'obiettivo fondamentale è sempre quello di mantenere quanto più possibile fedele all'originale l'immagine acquisita, indipendentemente dal dispositivo che la sta processando: scanner, monitor o stampante

Esistono, infatti, considerevoli limiti riguardo ai colori e altre caratteristiche che ogni periferica è in grado di produrre e visualizzare. Lo spazio colore dei diversi dispositivi è quasi sempre diverso e questo comporta il rischio di produrre differenti output

finali. Per comprenderne le differenze basta analizzare e confrontare lo spazio dei colori CIE[20] dell'occhio umano, di monitor LCD e di una stampante:

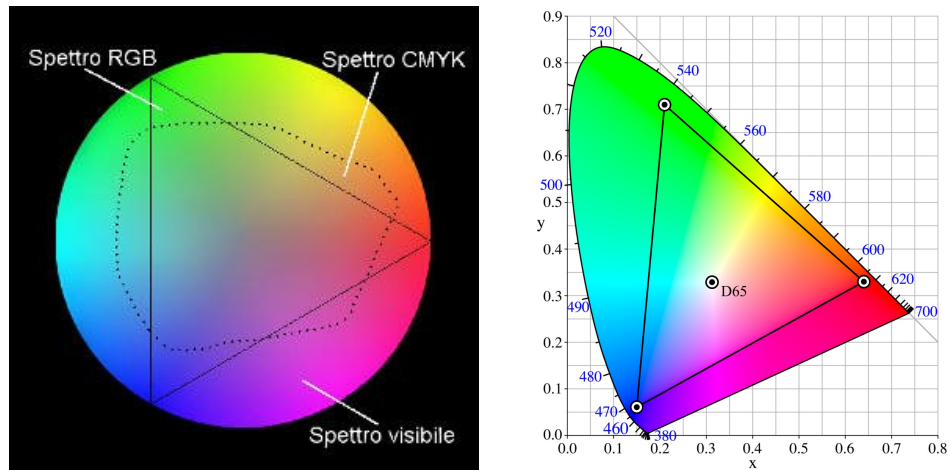



Figura 10.16: un confronto tra lo spazio colore visibile, RGB e CMYK

Esistono programmi specifici per la gestione del colore, che permettono di regolare la riproduzione del colore all'interno del flusso produttivo gestendo il passaggio da una gamma ad un'altra. Il problema della gestione dello spazio di colore è complesso e molte case di produzione di prodotti informatici forniscono linee guida per la calibrazione dei propri dispositivi, generalmente corredati da prodotti software inclusi nel pacchetto. In alternativa, sono disponibili in commercio applicazioni di terze parti molto valide.

Il processo di calibrazione comprende due fasi: la regolazione dei livelli di esposizione, messa a fuoco, mezzitoni,

[20] L'occhio umano presenta una diversa sensibilità alla radiazione a seconda della lunghezza d'onda. Nel 1931 la CIE ha definito le caratteristiche convenzionali di un occhio "medio" e le relative curve di visibilità. Questa definizione ha consentito di individuare il valore massimo di visibilità relativa per radiazioni di lunghezza d'onda pari a 555 nm per la visione fotopica e di 507 nm per quella scotopica.

ombreggiature di ciascun dispositivo di acquisizione ottica (scanner) e l'uniformazione delle caratteristiche di riproduzione del colore di tutti i dispositivi nel processo produttivo (monitor e stampanti).

10.1.1 Calibrazione del monitor

Esistono sia strumenti software che strumenti hardware per ottenere una corretta gamma di riproduzione del colore. Quasi tutti gli schermi possiedono alcuni profili di bilanciamento automatico preimpostati. Calibrando il monitor, in realtà si modifica l'intero sistema di visualizzazione del computer, comprese le importazioni della scheda video.

I fattori che influenzano la produzione del colore, sia in un monitor CRT (tubo a raggi catodici) che in un monitor LCD (a cristalli liquidi) sono:

1. **Le coordinate colorimetriche XYZ dei tre colori primari R, G e B alla massima intensità.** Queste impostazioni non sono modificabili. Ogni monitor ha tre colori primari (rosso, verde e blu) e le loro coordinate colorimetriche non sono controllabili e purtroppo soggette a modificarsi con il tempo e l'utilizzo nei monitor CRT.

2. **Le coordinate colorimetriche XYZ del bianco**. Nei monitor CRT il bianco è modificabile operando sull'hardware, con i cosiddetti cannoni elettronici o manualmente con i comandi OSD (*On Screen Display*), oppure via software mediante DDC (*Display Data Channel*). I monitor LCD possiedono un bianco nativo di 6500 K, che non è possibile modificare via hardware e software.

3. **I valori gamma dei tre colori primari R, G e B** (che spesso si considerano uguali). Sono modificabili esclusivamente agendo sulla LUT (*Look Up Table*) della scheda video.

Il primo passo per la calibrazione del monitor è regolare i valori di luminosità e contrasto nei CRT e di sola luminosità negli LCD (la luminosità degli LCD corrisponde al contrasto dei CRT). Solitamente si regola al massimo contrasto negli LCD e a metà luminosità nei CRT. Successivamente, attraverso un pannello OSD (*On Screen Display*) del monitor si regolano i valori gamma e del bianco in determinate posizioni; si consigliano i valori rispettivi di 2.2 e D65. Dopo che il monitor è rimasto acceso per almeno mezz'ora, si misurano i valori: con il programma che crea il profilo, che può servirsi di uno strumento di misura (colorimetro o spettrofotometro), oppure basarsi sulla valutazione dell'occhio

dell'utente. Entrambe le modalità determinano il punto bianco e il gamma del monitor: la prima li misura, la seconda presenta all'utente delle strutture di punti e linee e, ponendogli domande, cercano di determinare i valori da prendere come riferimento. L'utente specifica infine quale punto bianco e gamma desidera.

QuickTime™ e un
decompressore BMP
sono necessari per visualizzare quest'immagine.

Figura 10.17: schermata di calibrazione gamma di un monitor a 2.2. Fonte: http://www.ballisticpublishing.com

Il monitor CRT va calibrato periodicamente, ogni circa 300 ore di accensione; per quello LCD l’operazione può essere ripetuta con intervalli più lunghi, ad esempio ogni 6 mesi.

C'è inoltre da aggiungere che la gestione del colori avviene anche a livello di sistema operativo. Sia Macintosh che Windows adottano i profili comuni standard dell'ICC, anche se a livello empirico il bilanciamento gamma che si adotta sui monitor che gestiscono i due sistemi operativi è leggermente diverso. Questo avviene perchè i due sistemi operativi adottano due versioni di software diverso: Mac OSX utilizza *ColorSync 4.0*, Windows adotta *Windows Color System* (WCS).

10.1.2 Calibrazione di scanner e dispositivi di acquisizione ottica

Esistono in commercio delle tabelle di test formate da quadrati di diversi colori, denominate *colorchecker*, stampate su carta fotografica con la massima cura e testate con spettrofotometri professionali. Tramite un software di calibrazione si sottopone a un esame cromatico lo scanner: inserita la matrice di colori sul piano di lettura, il software di calibrazione analizza il comportamento del dispositivo. Confrontando i valori letti dallo scanner con i dati reali della matrice, il software di calibrazione è in grado di costruire il profilo cromatico dello scanner effettuando una sottrazione tra le due tabelle.

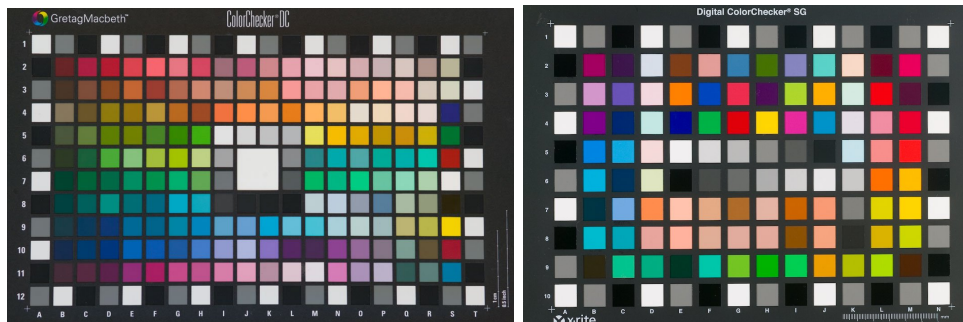



Figura 10.18: una *colorchecker* GretagMacbeth DC e una X-write

Al fine di risolvere il problema della fedeltà di riproduzione del colore nei diversi dispositivi digitali, nel 1993 è stato definito un profilo standard dall'*International Color Consortium* (ICC), un organismo fondato dai principali produttori internazionali di hardware e software, tra i quali compaiono Adobe, Agfa, Apple, Kodak, Microsoft, Silicon Graphics e Sun. Lo standard ICC prevede che lo stesso profilo possa essere utilizzato sia quando è l'origine che quando è la destinazione della conversione di colore, quindi contiene sia una tabella che va da coordinate di periferica a coordinate assolute (usata quando il profilo è l'origine della conversione), sia le tabelle che vanno da coordinate assolute a coordinate di periferica (usate quando il profilo è la destinazione della conversione). I profili ICC possono essere opportunamente creati oppure, più di frequente, distribuiti insieme al dispositivo e implementati in fase di configurazione, garantendo così la sicurezza di una opportuna interconversione delle coordinate cromatiche.

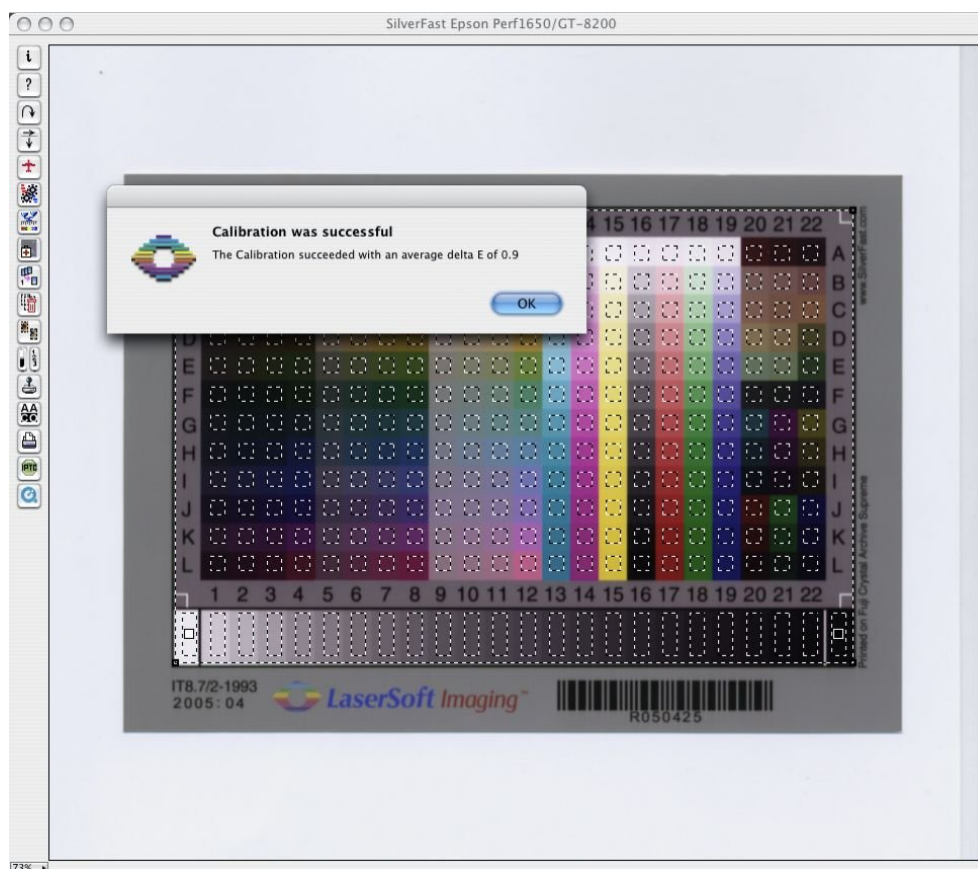



Figura 10.19: schermata di una calibrazione ottenuta con successo su uno scanner Epson con il software di calibrazione SilverFast Ai della casa di produzione tedesca LaserSoft Imaging

10.1.3 Ambiente e illuminazione

Benché non appartengano né al campo della strumentazione hardware né dell'architettura software, è bene mettere in pratica preliminarmente degli accorgimenti sull'ambiente entro il quale si svolgerà l'attività di scansione. Il luogo in cui si svolgerà la digitalizzazione dovrà avere una luce costante nell'arco della giornata, preferibilmente di tonalità fredda, e una colorazione

uniforme delle pareti. È preferibile collocare i dispositivi su un piano di lavoro di colore uniforme e neutro, che eviti riflessi di luce.

È raccomandabile, inoltre, eseguire frequentemente la pulizia dei ripiani degli scanner, per evitare la presenza di polvere o macchie che possono influire negativamente sulla qualità finale dell'immagine.

10.2 Posizionamento dell'originale

Il documento originale deve essere opportunamente posizionato per garantire la scansione corretta dell'area. Nel caso di scanner piano, la pagina da sottoporre a scansione va adagiata al centro del piano dello scanner, la zona con luce più uniforme, con la facciata da acquisire rivolta verso il basso, avendo cura di chiudere il coperchio. Con uno scanner planetario il volume va posizionato a faccia in su e centrato con precisione rispetto al piano di ripresa, che spesso riproduce una griglia per facilitare questa operazione. Se si intende procedere con la scansione di pellicole o diapositive, queste andranno invece inserite nell'apposito supporto fornito con lo scanner, in modo che il lato lucido sia rivolto verso il piano di copiatura quando la sorgente luminosa è in posizione (il lato dell'emulsione non è sensibile). Con uno scanner a inserimento fogli è sufficiente posizionare i documenti nell'apposito vassoio, mentre

nel caso di uno scanner a tamburo è necessario sistemare l'originale con molta precisione sul cilindro, trattenendolo con appositi liquidi o nastri adesivi.

10.3 Preparazione alla scansione e anteprima

L'acquisizione di un'immagine con un dispositivo ottico può essere effettuata in due modi: avviando il software fornito con lo scanner o, attraverso l'interfaccia *twain*, utilizzando la funzione "importa" di qualsiasi programma di fotoritocco o elaborazione di immagini. In entrambi i casi, il programma che gestisce la scansione richiede che siano indicati alcuni parametri per l'acquisizione: la tipologia dell'originale, le modalità di acquisizione (profondità di colore e intervallo dinamico), la destinazione (salvataggio o apertura automatica all'interno di un software di elaborazione) e, infine, i parametri di risoluzione della scansione che determinano le dimensioni finali. Si è già ampiamente discusso delle specifiche da considerare in relazione alla definizione dei livelli qualitativi.

Definite le impostazioni dei parametri di acquisizione, si può procedere con la pre-scansione. Questa restituisce un'anteprima di scansione (a bassa risoluzione) tramite la quale è possibile avere una rappresentazione sommaria del risultato dell'immagine che si intende digitalizzare. È possibile, ad esempio, selezionare l'area di

acquisizione occupata dal documento, per evitare di campionare uno spazio di dimensioni superiori a quelle occupato della pagina da scansire e ritrovare nel file digitalizzato un margine vuoto da ritagliare successivamente. Oltre al ritaglio e al ridimensionamento dell'anteprima è possibile regolare la gamma e la tonalità e correggere le dominanti di colore.

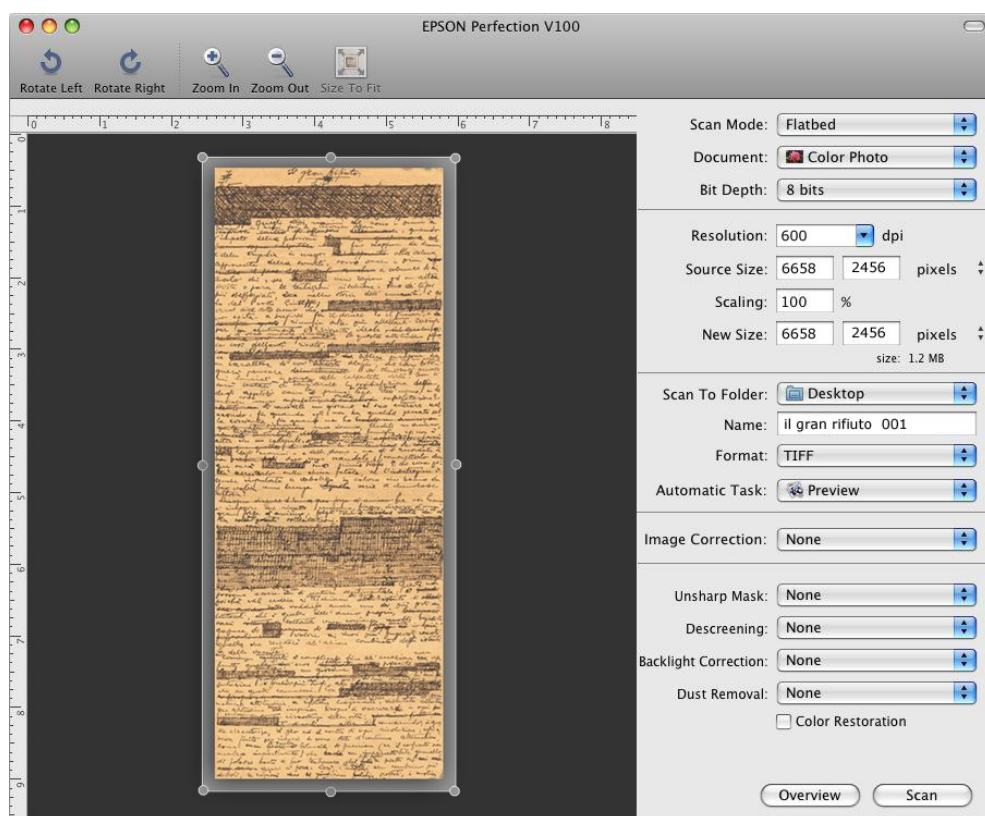



Figura 10.20: anteprima di scansione della prima pagina del manoscritto *Il gran rifiuto* di Federico De Roberto.

Di solito i software di preparazione preliminare della scansione non permettono grandi interventi sull'immagine, ma

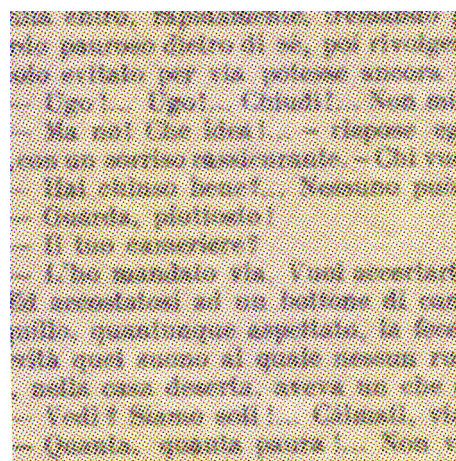
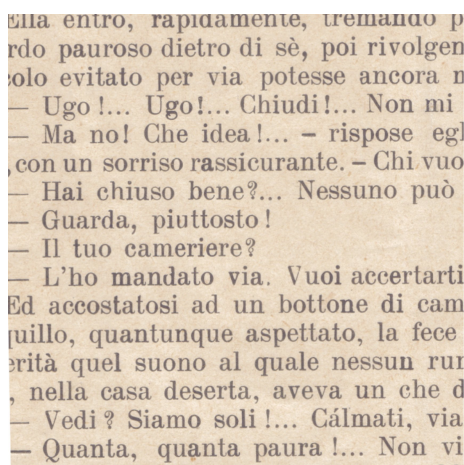
presentano delle funzionalità essenziali per il mantenimento della fedeltà del file digitalizzato rispetto all'originale. Ad esempio, per digitalizzare un documento stampato, è necessario eseguire una rimozione della retinatura durante la scansione; in caso contrario si otterrebbero difetti di interferenza visiva (effetto *moiré*[21]) e un'immagine digitale visibilmente composta da numerosi puntini, piuttosto difficili da eliminare in un momento successivo. Quasi tutte le interfacce di scanner includono una funzione di rimozione retinatura chiamata *dithering removal* o *descreening*. In genere, per questa funzione, sono disponibili diverse opzioni a seconda del documento che si intende digitalizzare: foto, carta stampata, rivista a colori e così via. È sufficiente selezionare quella che si ritiene più idonea. Va ricordato un eccesso nella rimozione della retinatura tende a rendere l'immagine finale leggermente sfocata.

[21] Con *effetto moiré* si indica un figura di interferenza, creata ad esempio da due griglie sovrapposta ad un certo angolo, o anche da griglie parallele con maglie distanziate in modo leggermente diverso.

Figura 10.21 e Figura 10.22: dettaglio della prima pagina delle bozze di stampa de *Il cane della favola* di Federico De Roberto: un confronto tra l'immagine acquisita con l'operazione di deretinatura e senza.

10.4 Scansione finale

Dopo aver impostato tutti i parametri e verificato il risultato attraverso l'anteprima, è necessario completare il processo dando avvio alla scansione vera e propria, che darà origine al file finale, da salvare nel formato di output più idoneo. La scansione può richiedere tempi più o meno lunghi, a seconda della tipologia di dispositivo impiegato e alle specifiche impostate. Ad ogni modo, questa fase non richiede nessun intervento da parte dell'operatore. La fase successiva sarà quella di editing del file digitale ottenuto.

11. Editing digitale

L'elaborazione, o editing, digitale delle immagini è una operazione che comporta l'utilizzo di algoritmi per effettuare delle specifiche modifiche sulle stesse. Tali algoritmi agiscono partendo dal presupposto che le immagini raster sono memorizzate in forma di matrice di pixel, ognuno dei quali contiene un dei valori per identificare il colore e la luminosità del punto. I programmi possono cambiare i valori dei pixel per modificare l'immagine, utilizzando diversi modelli; solitamente si tratta di valutare un gruppo di pixel adiacenti e, in base a questi, stabilire i valore di colore e luminosità più adatti. Applicando trasformazioni numeriche, gli algoritmi restituiscono una immagine modificata, oppure un dato numerico o tabellare rappresentativo di una particolare caratteristica dell'immagine in input.

L'immagine ottenuta attraverso la scansione può presentare alcuni elementi da trattare con un programma di *editing* grafico.

11.1 Software per l’elaborazione delle immagini

Esiste un gran numero di software di editing per immagini, sia proprietari che *open source*. Il top di gamma in campo di software proprietario è Adobe Photoshop[22], giunto alla sua dodicesima versione[23] inclusa nella Creative Suite 5 della casa di produzione Adobe: è il più completo e offre un numero enorme di funzionalità. Una versione più leggera, dedicata all’editing fotografico e con una nuova interfaccia, è stata rilasciata nel 2007 con il nome di Adobe Lightroom. Lightroom è giunto di recente alla versione 3.0. è consigliabile per questo tipo di lavoro perché la sua interfaccia consente non solo un editing più che sufficiente per i file presi in considerazione, ma soprattutto perché la sua interfaccia consente di organizzare e catalogare meglio grandi volumi di immagini digitali, con l’utilissima funzionalità di tenerne sempre sott’occhio l’anteprima. Inoltre, uno dei vantaggi più rilevanti consiste nell’editing non distruttivo. Gli interventi eseguiti sui file all’interno di Photoshop Lightroom non vanno mai ad alterare le immagini originali, ma vengono memorizzati nel database chiamato *Catalog*. Questo aspetto è essenziale sia per la qualità del risultato sia per la modularità e reversibilità delle proprie regolazioni. È

[22] Adobe Photoshop è un'applicazione informatica prodotta dalla Adobe Systems Incorporated specializzata nell'elaborazione di fotografie e, più in generale, di immagini digitali. La prima versione di Photoshop risale al 1990 per opera dei fratelli Knoll, figli di un fotografo e autori del programma per agevolare il lavoro del padre. Di Adobe Photoshop esiste anche una versione semplificata e più economica, chiamata Photoshop Elements.

[23] La versione attualmente disponibile in italiano è la CS5. Si tratta della versione 12, sebbene non venga più indicata con il numero come si era fatto sino alla 7.0: questo perché la nuova sigla CS ricorda che il software è parte integrante della Adobe Creative Suite, composta nella versione standard da Photoshop, Illustrator, InDesign, Version Cue, Bridge, Stock Photos, a cui si aggiungono GoLive e Acrobat Professional nella versione Premium. Tutti i componenti della Creative Suite sono comunque disponibili separatamente.

potenzialmente possibile effettuare infinite correzioni all'immagine che vengono applicate, in un unico passaggio e in modo irreversibile, soltanto in fase di esportazione, con la creazione di una copia dell'originale. Tutti i programmi che lavorano con immagini digitali costituite da pixel utilizzano generalmente invece un approccio cosiddetto "distruttivo". Non appena l'immagine è salvata e chiusa, qualunque intervento diventa permanente e l'immagine perde definitivamente alcune delle sue caratteristiche originali. Tutte le operazioni eseguite all'interno di Photoshop Lightroom sono memorizzate all'interno del database del programma in modo automatico. Questo significa che non occorre mai salvare il file. Tra l'altro, anche a seguito di una chiusura inaspettata del sistema operativo o di una interruzione di corrente, tutte le correzioni svolte sono memorizzate nel database fino all'ultima operazione.

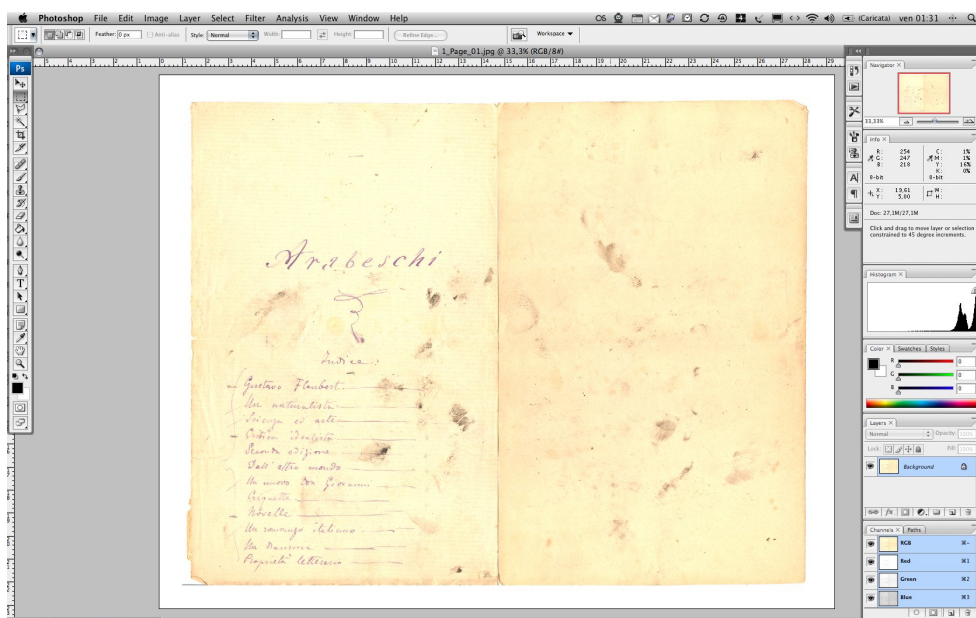



Figura 11.23: interfaccia utente di Adobe Photoshop

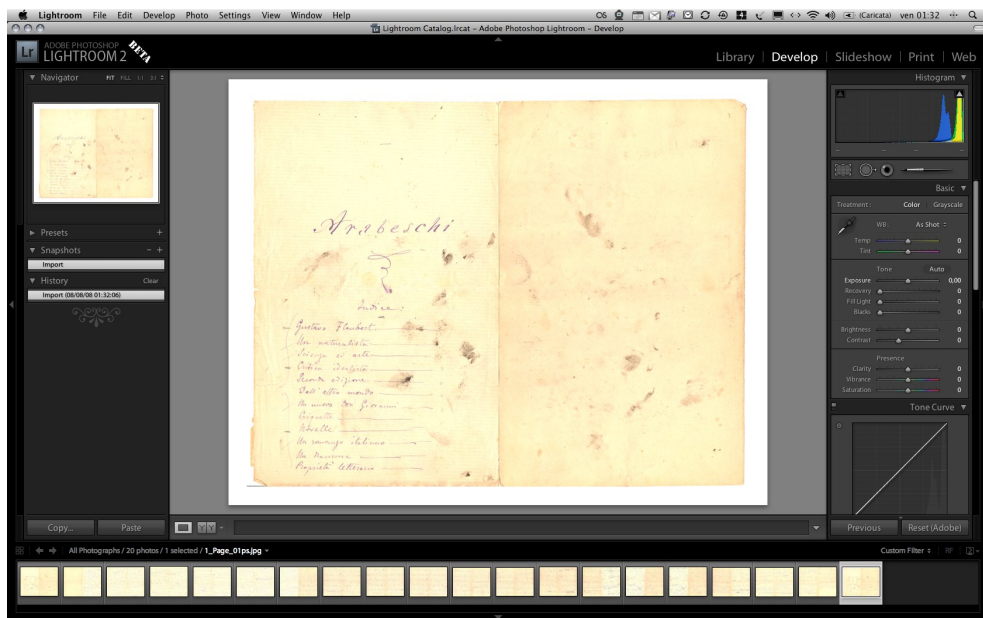



Figura 11.24: interfaccia utente di Adobe Photoshop Lightroom. Mantiene il layout in ciascuno dei moduli che costituiscono il programma (library, develop, slideshow, print e web) e si compone di 5 parti:
- L'area operativa centrale, nella quale si trova la o le immagini su cui si sta lavorando.
- La barra superiore, che contiene i links per passare da un modulo all'altro.
- La barra inferiore, occupata dalla filmstrip nella quale sono visualizzate le immagini dello shoot o della collection attivi.
- Il pannello di sinistra nel quale sono presenti impstazioni predefinite dette presets.
- Il pannello di destra nel quale sono presenti tutti i comandi disponibili per il modulo attivo.

Altri software per l'elaborazione grafica molto validi sono Corel Draw photopaint, Jasc Paint Shop Pro, Apple Aperture, Ulead Photoimpact. Nel caso di un progetto realizzato con software libero, Gimp per Linux Ubuntu si rivela altrettanto efficiente.

11.2 Operazioni di *editing*

Le principali operazioni di editing, necessarie alla trattazione delle immagini in seguito all'acquisizione, sono il ritaglio, il

raddrizzamento, la regolazione dei valori di luminosità e contrasto, della nitidezza e il bilanciamento dei colori.

11.2.1 Ritaglio (crop)

Dall'immagine scansita vanno opportunamente rimossi elementi e spazi superflui prodotti in fase di scansione. È opportuno, quindi, eliminare i bordi e i margini acquisiti creati dallo scanner attorno alla pagina sulle zone non occupate dal foglio.

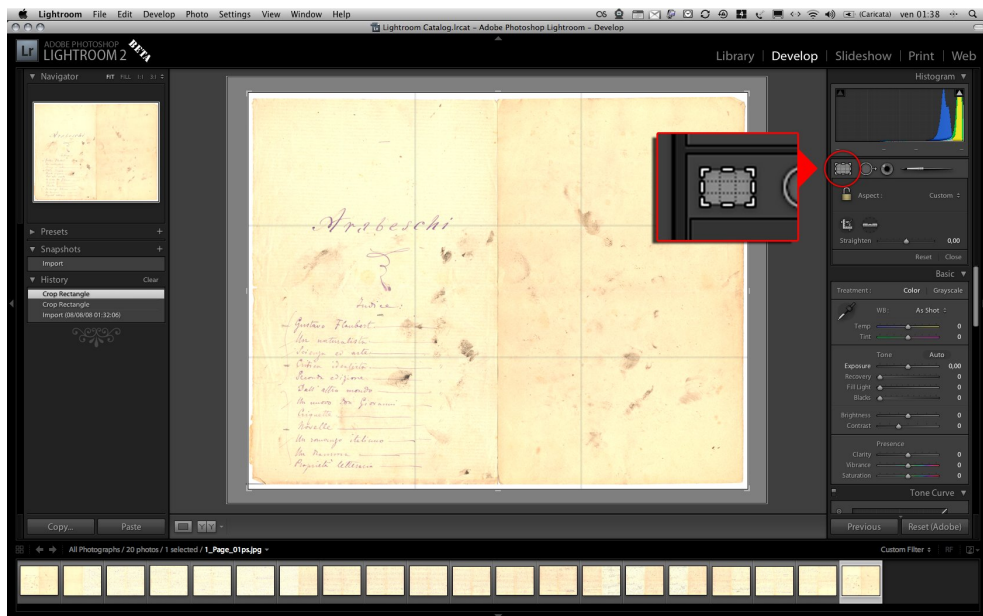



Figura 11.25: esempio di una operazione di ritaglio con Adobe Photoshop Lightroom sul frontespizio del manoscritto *Arabeschi* di Federico De Roberto. Dal pannello *develop* selezionare lo strumento *crop overlay* (shortcut da tastiera: "R"). Appare un bordo dotato di "maniglie" che circonda l'intera foto. Se la casella *Constrain Aspect Ratio* non è attivata le maniglie possono essere trascinate indipendentemente effettuando il ritaglio; se invece è attiva, l'area di ritaglio mantiene le proporzioni definite nel menù a tendina sulla destra. Una volta definita l'area di ritaglio, questa può essere spostata trascinando l'immagine. Digitando sulla tastiera il tasto "L" una volta si otterrà l'abbassamento della luminosità intorno all'area di lavoro, mentre premendolo una seconda volta le luci si spegneranno completamente, lasciando illuminata solo l'area ritagliata. In questo è possibile visualizzare l'immagine in un'area completamente nera, e valutare il risultato di questa operazione in maniera più precisa.

11.2.2 Raddrizzamento (straighten)

È molto probabile che l'immagine scansita appaia non perfettamente dritta, ma ruotata di qualche grado o decimo di grado. È bene raddrizzarla servendosi della una griglia visualizzata dal software che consenta un allineamento preciso. È sufficiente prendere in considerazione i margini superiori o inferiori e laterali della pagina, usandoli come linee guida, e disporli parallelamente ai segmenti della griglia virtuale.

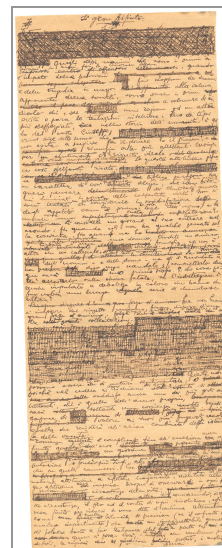



Figura 11.26: Prima pagina del manoscritto *Il gran rifiuto* di Federico De Roberto. L'immagine non risulta perfettamente dritta, ma inclinata di qualche grado

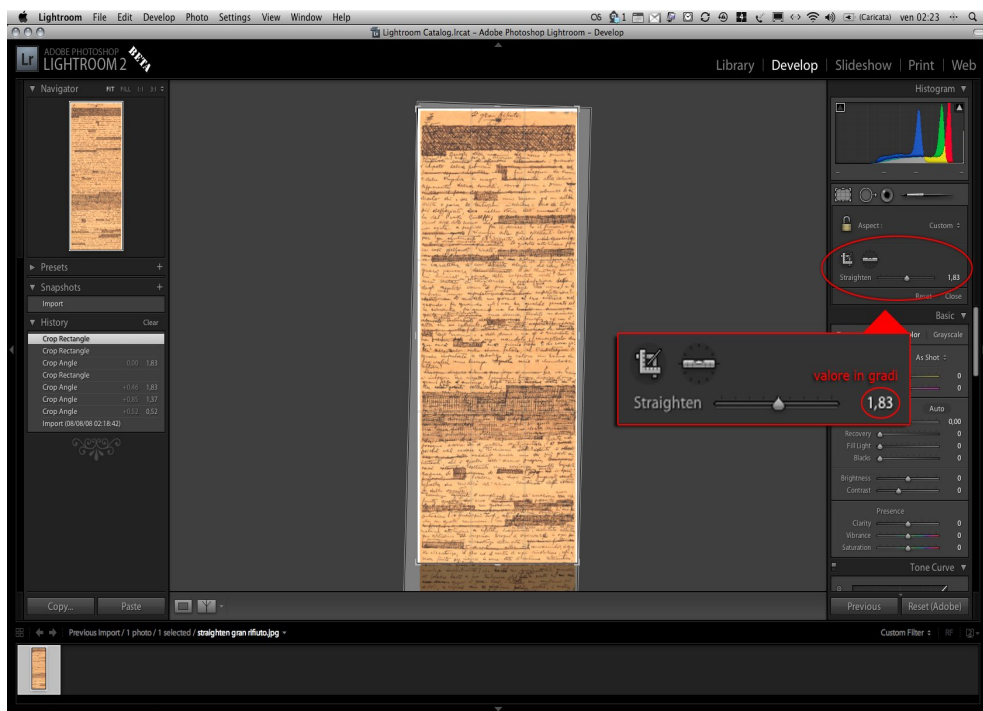



Figura 11.27: esempio di una operazione di raddrizzamento con Adobe Photoshop Lightroom sulla prima pagina del manoscritto *Il gran rifiuto* di Federico De Roberto. Dal pannello *develop* selezionare lo strumento *crop overlay* (shortcut da tastiera: "R"). Appare un bordo dotato di "maniglie" che circonda l'intera foto. È possibile ruotare l'immagine attraverso due diversi metodi: trascinando le maniglie ai quattro angoli dell'immagine o spostando il cursore a destra o a sinistra sulla barra all'interno del pannello per ruotare rispettivamente in senso orario o antiorario. L'angolo di rotazione, espresso in gradi, è visualizzato all'interno dello stesso pannello.

11.2.3 Luminosità e contrasto (*brightness/contrast*)

Le immagini acquisite sono comunemente generate dallo scanner con un buon grado di precisione nelle tonalità, ma contrasto e luminosità solitamente necessitano di un intervento. Solitamente la regolazione della luminosità e del contrasto è utilizzata per la correzione di immagini che sono troppo scure o troppo chiare. Un contrasto carente o una luminosità insufficiente dell'immagine possono disturbare l'effetto globale o modificare

sostanzialmente l'immagine digitale dalla sua versione originale, tanto da richiedere una successiva elaborazione.

La luminosità consente di regolare i mezzitoni dell'immagine. Si può lavorare sulla luminosità per mostrare più dettagli nelle zone d'ombra oppure ridurre le zone fortemente illuminate. Regolando la luminosità, ciascuna componente dell'immagine diventa più chiara o più scura.

Il contrasto è il rapporto o differenza tra il valore più alto (punto più luminoso) e il valore più basso (punto più scuro) nell'immagine. Incrementando tale rapporto, i valori più luminosi tendono al valore massimo (255) e i valori più scuri tendono al valore minimo (0). I valori intermedi non cambiano. In altre parole, aumentando il contrasto, le luci (zone più chiare) diventano più chiare e le ombre (zone più scure) si scuriscono. Viceversa, diminuendo il contrasto, la differenza tra chiaro e scuro si riduce, determinando una predominanza di toni intermedi.

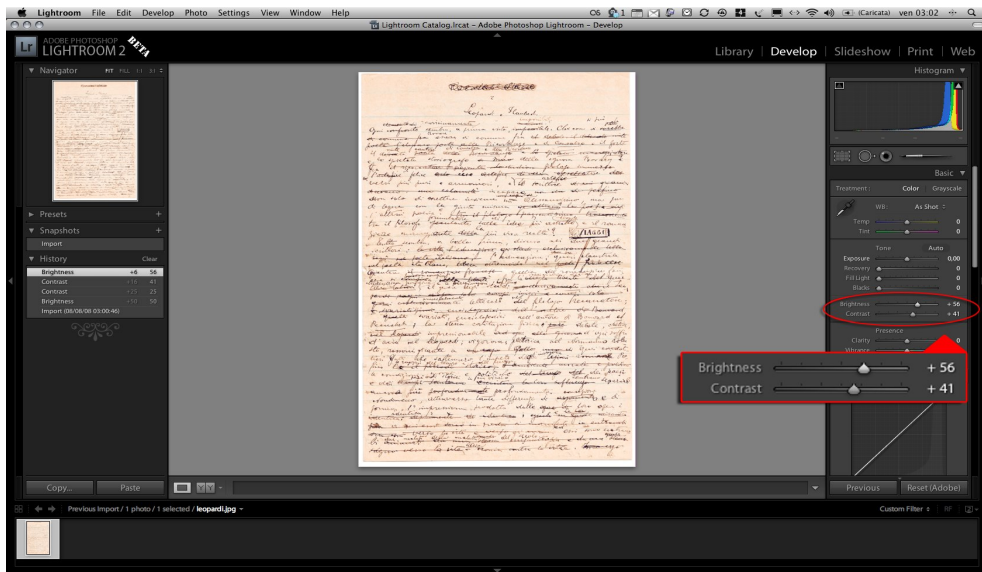



Figura 11.28: esempio di una operazione di regolazione di luminosità e contrasto con Adobe Photoshop Lightroom sulla prima pagina del saggio manoscritto *Leopardi e Flaubert* di Federico De Roberto. Dal modulo *develop* selezionare il pannello degli strumenti *basic* e regolare le impostazioni di *brightness* (luminosità) e *contrast* (contrasti), spostandosi con Il cursore a destra o a sinistra rispettivamente per aumentare o ridurre i valori.

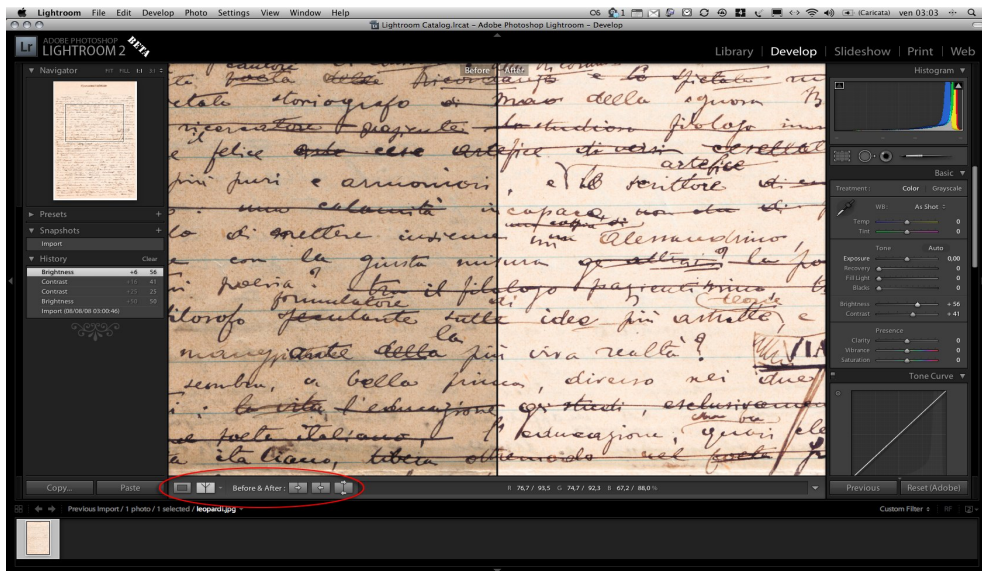



Figura 11.29: esempio di una operazione di regolazione di luminosità e contrasto con Adobe Photoshop Lightroom sulla prima pagina del saggio manoscritto *Leopardi e Flaubert* di Federico De Roberto: un confronto tra l'immagine prima della correzione e dopo (dettaglio, scala 1:1).

11.2.4 Migliorare l’OCR

Per migliorare la successiva fase di riconoscimento ottico del testo è consigliabile convertire l’immagine in scala di grigi e aumentare considerevolmente i valori di luminosità e contrasto. È un passaggio facoltativo; tuttavia può migliorare la prestazione dell'OCR (il cui risultato dovrà comunque essere sottoposto a controllo).

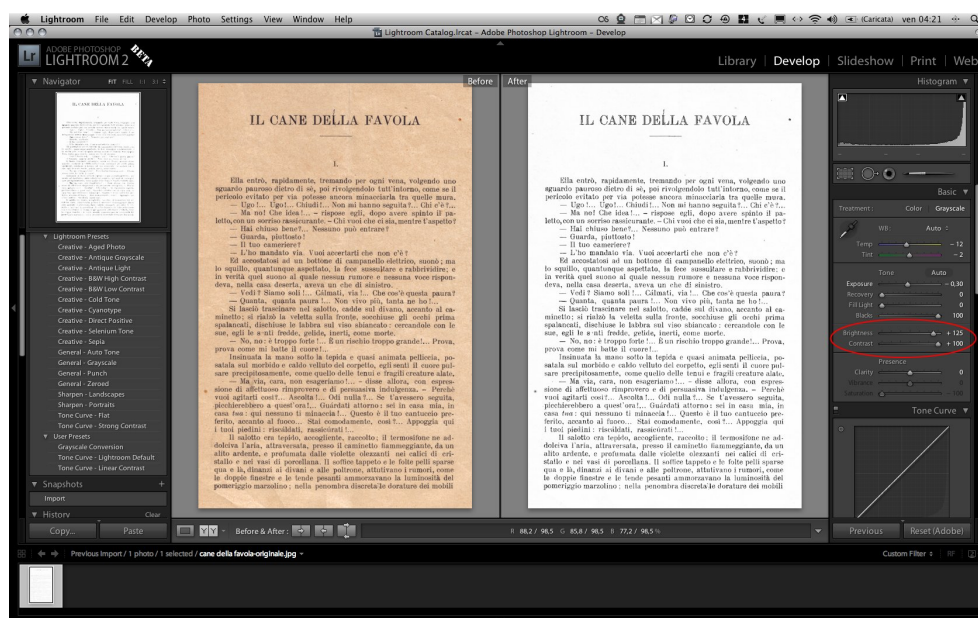



Figura 11.30: esempio di una operazione di regolazione di luminosità e contrasto in preparazione al riconoscimento caratteri con Adobe Photoshop Lightroom sulla prima pagina delle bozze di stampa de *Il cane della favola* di Federico De Roberto: un confronto tra l’immagine prima della correzione e dopo.

È possibile verificare questa ipotesi con un esperimento. Sottoponiamo la scansione di una pagina al riconoscimento ottico del testo e valutiamone i risultati in termini di tempo di scansione del programma e parole incerte restituite:

	Tempo di editing	Tempo di scansione	Parole incerte
Immagine originale	0”	6,9”	10
Immagine ritoccata	120”	4,1”	5

Tabella 11.19: risultati di un esperimento di riconoscimento ottico dei caratteri prima e dopo l’editing digitale dell’immagine.

Dai risultati si può osservare come effettivamente il miglioramento del contrasto e della luminosità influenzi il numero di parole incerte restituite. Con una semplice operazione di editing è possibile dimezzare il successivo lavoro di controllo necessario: un risparmio di tempo nella successiva analisi degli errori a fronte di un minimo tempo speso per l'editing.

11.2.5 Nitidezza e dettaglio (sharpening)

Mediante l’*editing* dello *sharpening* si cerca di recuperare parte della nitidezza persa durante il processo di campionamento dell’immagine. Ovviamente non si possono creare dettagli che non esistono, ma i risultati sono inaspettatamente buoni. È possibile migliorare significativamente la leggibilità e la qualità della visualizzazione dei tratti tipografici, poiché aumenta il contrasto lungo i bordi, ma la conseguenza è un aumento di rumore nelle zone di colore uniforme. È quindi necessario impostare un valore che rappresenti un buon compromesso.

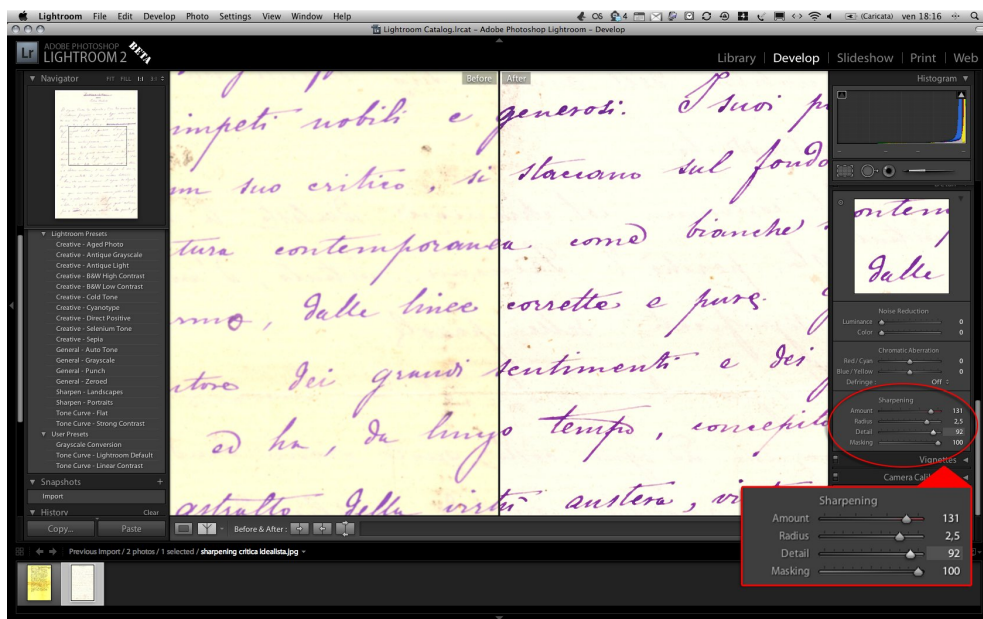



Figura 11.31: esempio di una operazione di regolazione di dettaglio con Adobe Photoshop Lightroom sulla prima pagina della *Critica idealista* di Federico De Roberto: un confronto tra l'immagine prima della correzione e dopo (dettaglio, scala 1:1). Dal modulo *develop* selezionare il pannello degli strumenti *detail* e regolare le impostazioni di *sharpening* spostandosi con il cursore a destra o a sinistra rispettivamente per aumentare o ridurre i valori.

11.2.6 Bilanciamento dei colori (*color balance*)

Durante le fasi finali d'elaborazione dell'immagine digitale ci si trova ad affrontare il problema di rappresentare i colori dell'immagine nella maniera più fedele possibile all'originale. Talvolta le immagini scansite possono avere delle indesiderate dominanti di colore. È bene ricordare che per ottenere un effetto realistico durante il processo di elaborazione rispetto all'originale è necessario lavorare su uno schermo calibrato opportunamente.

Il metodo migliore per effettuare il bilanciamento del colore è agire su intervenire un singolo colore per volta nell'ambito di una immagine. Lo spazio HLS (*Hue Luminance Saturation*) consente di

intervenire su ogni intervallo di colore modificandone rispettivamente la tonalità (*hue*), che regola il range di tinta di ogni singolo colore, la luminanza (*luminance*), che regola la luminosità del colore, e la saturazione (*saturation*), che regola l'intensità di ogni singolo colore. È possibile, per esempio, correggere la dominanza del giallo e dell'arancio, ottenuta dalla scansione di un foglio ingiallito dal tempo, abbassando la saturazione dei colori specifici senza alterare tutti gli altri.

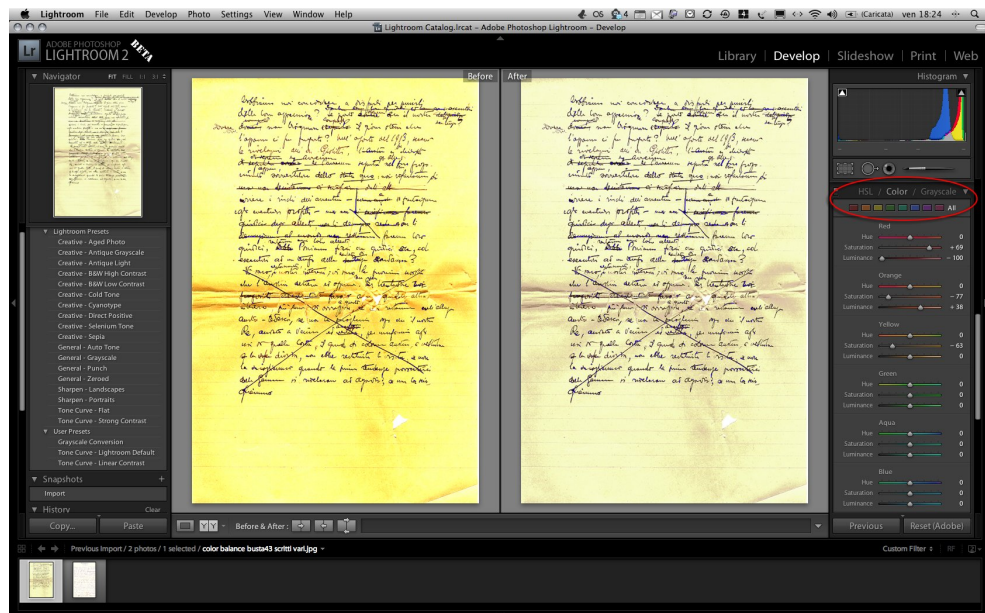



Figura 11.32: esempio di una operazione di bilanciamento del colore con Adobe Photoshop Lightroom sulla prima pagina di un insieme di *scritti vari* (busta 43) di Federico De Roberto: un confronto tra l'immagine prima della correzione e dopo (dettaglio, scala 1:1). Dal modulo *develop* selezionare il pannello *HLS* e regolare, per ogni singolo colore da bilanciare le impostazioni di *Hue* (tonalità), *Saturation* (saturazione) e *Luminance* (luminanza). Il Pannello *Color* permette di svolgere le stesse funzioni di Pannello HSL raggruppate per colori.

11.3 Riconoscimento dei caratteri e del testo: OCR e ICR

La conversione delle immagini acquisite in testo modificabile è una fase che richiede un impegno considerevole di risorse, sia per impiego di tempo che di personale; è quindi una questione strettamente correlata anche alle risorse economiche. L'operazione di riconoscimento dei caratteri digitali è eseguita da programmi OCR, ma è comunque necessario l'intervento umano per un controllo finale.

I software di riconoscimento ottico dei caratteri sono ormai molto evoluti e affidabili. Il riconoscimento esatto di un testo scritto in alfabeto latino è considerato un problema risolto quasi in maniera definitiva, raggiungendo livelli di riconoscimento dei caratteri superiori al 99%. Il riconoscimento della scrittura a mano libera e il riconoscimento degli alfabeti non latini è un problema che tuttora non ha trovato delle soluzioni realmente soddisfacenti, ed è tuttora oggetto di studi e ricerche; lo stesso si può dire dei sistemi di scrittura a mano libera o in corsivo.

Le immagini da sottoporre alla scansione ottica per la lettura dei caratteri devono avere delle caratteristiche specifiche: i software OCR lavorano meglio su immagini fortemente contrastate, come accertato precedentemente, ma hanno soprattutto bisogno di un range di risoluzione adeguato, che si attesta tra un minimo di 72 dpi e un massimo di 600 dpi. Oltre questi limiti sarà difficile ottenere

buoni risultati: con bassi livelli di risoluzione si ha una cattiva prestazione di riconoscimento del testo; con risoluzioni alte, invece, si ha una complessità di calcolo molto elevata e un dispendio eccessivo in termini di tempo.

Il formato di output generato dal software dipende dalle impostazioni dell'utente. È possibile scegliere formati che mantengono l'informazione di layout della pagina, mantenendone inalterata l'impaginazione e la formattazione, oppure estrarre esclusivamente l'informazione testuale.

I software OCR sono molto diffusi. Ne esistono diverse versioni sia proprietrarie (Omnipage, TextBridge, ABBYY Finereader, Gocr, WiseImage, Readiris, Nuance) che *open source* (GNU Ocrad, Tessereact).

Il sistema ICR (*Intelligent Character Recognition*) è un'infrastruttura tecnologica ad apprendimento probabilistico, basata su una piattaforma neurale, per una ricognizione dei caratteri basata sugli studi di intelligenza artificiale. L'ICR attua una ricognizione ottica come quella dell'OCR ma, a differenza di quest'ultimo, la sua capacità di riconoscimento è molto simile all'intelligenza umana: applica soluzioni euristiche ai problemi.

Dopo il processo di apprendimento, l'ICR riconosce un carattere o un gruppo di caratteri individuando tra i logotipi da esso appresi quello che più probabilmente si avvicina all'originale.

In questo modo, l'ICR può imparare a riconoscere, in maniera sempre più raffinata e veloce, anche caratteri non usuali e non standardizzati, come i caratteri a stampa più antichi o quelli manoscritti.

12. Metadati

Un metadato, dal greco *meta-* "oltre" e dal latino *datum* (plurale *data*), "informazione", è una informazione che descrive un insieme di dati. In accordo con il suo significato etimologico, secondo la definizione della studiosa Anne Gilliland-Swetland, un metadato è «the sum total of what one can say about any *information object*». Si precisa inoltre:

> Cultural heritage and information professionals such as museum registrars, library catalogers, and archival processors are increasingly applying the term *metadata* to the value-added information that they create to arrange, describe, track and otherwise enhance access to information objects[24].

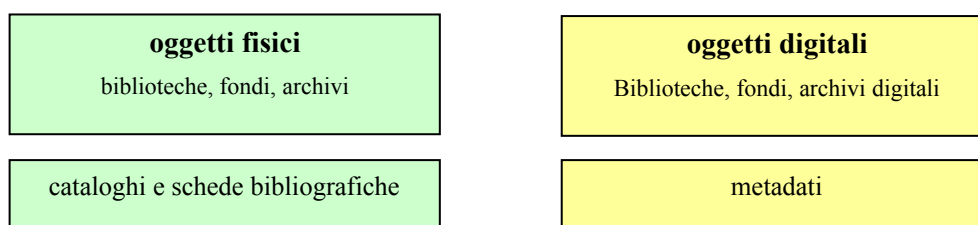
Tra le definizioni italiane, quella di Cristina Magliano dell'ICCU descrive i metadati come

> Informazioni strutturate che descrivono, specificano o localizzano una risorsa; i metadati permettono inoltre la sua ricerca, uso e organizzazione e sono la chiave per assicurarne la futura

[24] Gilliland-Swetland, *Setting the Stage* in *Introduction to Metadata*, <http://www.getty.edu/research/institute/standards/intrometadata/2_articles/index.html>.

accessibilità nel tempo[25].

Il termine *metadato* è oggi comunemente utilizzato per riferirsi all'ambito dell'informazione elettronica in rete: i metadati sono generalmente intesi come una riproduzione amplificata e migliorata in termini di prestazioni delle tradizionali pratiche di catalogazione bibliografica:

oggetti fisici biblioteche, fondi, archivi	**oggetti digitali** Biblioteche, fondi, archivi digitali
cataloghi e schede bibliografiche	metadati

Un set di metadati descrive, quindi, le risorse informative a cui si riferisce, per facilitarne l'accessibilità e la diffusione. Tutti i documenti sono indicizzati attraverso stringhe descrittive contenute in *record*, l'unità base di un database, che contiene un insieme di elementi, ciascuno dei quali possiede nome e tipo propri. Gli elementi di un record sono detti anche *campi*, e sono identificati da un nome specifico.

Per poter sfruttare con successo le potenzialità dei metadati, è indispensabile l'*information retrieval* (IR), l'insieme delle tecniche

---

[25] Cristina Magliano, *Lo standard nazionale dei metadati gestionali amministrativi*, in «Digitalia», n.0, Anno I, 2005, pag. 35.

utilizzate per il recupero dell'informazione in formato elettronico. I sistemi IR usano i linguaggi di interrogazione basati su comandi testuali.

Una ricerca di *information retrieval* ha come base di input un comando dell'utente, costituito da una o più parole chiave associate all'informazione richiesta. La *query*, l'interrogazione vera e propria rivolta al database, è messa in relazione con gli oggetti presenti nella banca dati. In output, il sistema fornisce un insieme di *record* che soddisfano le condizioni richieste.

12.1 Funzionalità dei metadati

I metadati possono essere distinti in tre categorie funzionali:

1. metadati descrittivi. La funzione di questa categoria di metadati consiste nell'identificazione e il recupero degli oggetti digitali. Le informazioni sono costituite da descrizioni uniformate dei documenti originali o nati in formato digitale, immagazzinate in genere nei database di *information retrieval* all'esterno degli archivi collegati agli oggetti digitali attraverso appositi collegamenti;

2. metadati amministrativi e gestionali. La funzione di questa categoria di metadati consiste nell'evidenziare le modalità di

archiviazione e manutenzione degli oggetti nel sistema di gestione dell'archivio digitale, per una corretta esecuzione delle attività pertinenti. Questi metadati, quindi, assumono un'importanza strategica ai fini della conservazione permanente degli oggetti digitali: i metadati aministrativi e gestionali possono documentare i processi tecnici associati alla conservazione permanente, fornire informazioni sulle condizioni e i diritti di accesso ai files, certificare l'autenticità e l'integrità del contenuto, documentare la catena di custodia degli oggetti e identificarli in maniera univoca. Comprendono inoltre metadati tecnici, che descrivono le caratteristiche tecniche della risorsa digitale; metadati per la conservazione relativi alla fonte e alla provenienza, che descrivono l'oggetto dal quale è derivata la risorsa digitale e la storia delle operazioni effettuate fin dalla sua creazione e infine metadati per la gestione dei diritti, che descrivono i diritti d'autore e di riproduzione, le restrizioni e le licenze che vincolano l'uso della risorsa;

3. metadati strutturali. La funzione di questa categoria di metadati consiste nel descrivere la struttura interna fisica o logica dei documenti (come ad esempio l'introduzione, i capitoli, le sezioni e l'indice di un libro) e le loro relazioni fra le varie parti degli oggetti digitali. Attraverso i metadati strutturali è possibile collegare le varie

componenti delle risorse, per un'adeguata e completa fruizione attraverso la mappatura di schemi di metadati diversi. Forniscono inoltre dati di identificazione e localizzazione del documento, come il codice identificativo, l'indirizzo del file sul server, l'archivio digitale di appartenenza e il suo indirizzo url.

Ogni comunità responsabile dell'organizzazione di differenti tipi di risorse ha sviluppato i propri standard per supportare diverse operazioni sui propri domini d'interesse. La comunità museale ha elaborato gli standard *Spectrum* e CDWA (*Categories for the Description of Works of Art*); la comunità archivistica ha sviluppato l'ISAD(G) (*General International Standard Archival Description*), l'ISAAR(CPF) (*International Standard Archival Autorithy Records for Corporate Bodies, Persons and Families*) e l'EAD (*Encoded Archival Description*) per i record d'archivio; la comunità bibliotecaria usa i formati della famiglia MARC (*MAchine Readable Cataloguing*) per la rappresentazione e lo scambio dei metadati bibliografici e ha definito anche standard descrittivi, le varie ISBD (*International Standard Bibliographic Description*). In altri ambiti sono stati definiti standard i metadati basati su linguaggi di marcatura quali lo *Standard Generalised Markup Language* (SGML) o l'*eXtensible Markup Language* (XML).

Il problema che si è manifestato negli ultimi anni ha interessato sempre più la ricerca di standard di metadati

interoperabili, muovendosi sempre più verso un modello che, rispettando le specificità e tenendo conto dei già profili esistenti, condividesse un set minimo di dati, cercando di utilizzare convenzioni e semantiche condivise ai fini della descrizione, ricerca e identificazione.

In Italia, nel muoversi in questa direzione, è stato costituito nel 2003 un gruppo di lavoro permanente: il comitato MAG.

12.2 Gli standard per i metadati

12.2.1 MAG

Dalla necessità di regolamentare i numerosi progetti in fase di avviamento in Italia per la creazione di set standard di metadati gestionali e amministrativi, l'Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche (ICCU), in qualità di responsabile della diffusione delle normative e degli standard bibliografici, ha creato nel 2003 un gruppo di lavoro permanente. Il comitato MAG, portando avanti le attività di studio sui modelli e gli ambiti di applicazione dei metadati per il patrimonio culturale, si è concentrato sugli sviluppi e la disseminazione nel campo dei Metadati Amministrativi Gestionali (MAG).

Gli obiettivi dichiarati dal comitato MAG sono:

- diffusione dello Schema MAG
- mantenimento ed evoluzione dello schema
- produzione di manuali e linee guida
- assistenza agli implementatori
- formazione e promozione
- rapporti con altri progetti e agenzie (progetti europei, *Dublin Core* ecc.)[26]

Il Gruppo ha collaborato con Biblioteca Digitale Italiana (BDI) su uno studio di fattibilità commissionato dal Ministero per i Beni e le Attività Culturali. Dall'indagine è stata messa in luce l'esigenza di documentare gli aspetti fisici della conservazione e della destinazione delle risorse librarie. Il modello analizzato è stato il modello di tipo logico-funzionale OAIS (*Open Archival Information System*). Data l'eterogeneità dei dati da documentare e gestire, si è messo in evidenza la necessità di un lavoro che coinvolgesse tutte le varie figure impegnate raccolta di tutti questi elementi deve essere il frutto di un lavoro coordinato dei vari attori coinvolti nelle fasi di creazione, distribuzione, analisi e accessibilità delle risorse. Da risultati dello studio è emerso che in Italia i primi progetti di digitalizzazione si sono limitati agli aspetti tecnici di riproduzione e

[26] Comitato MAG, *MAG Metadati Amministrativi e Gestionali. Manuale Utente*, a cura di Elena Pierazzo, versione: 2.0.1, ultimo aggiornamento 8 marzo 2006, <http://www.iccu.sbn.it/upload/documenti/manuale.html>.

scansione, senza documentare però le operazioni svolte, vanificando l'utilizzo a lungo termine delle risorse digitalizzate. Per un corretto futuro sviluppo è necessario disporre infatti non solo di standard relativi ai dati bibliografici, da sempre oggetto di analisi da parte delle biblioteche, ma anche di schemi di metadati amministrativi gestionali relativi a elementi quali le condizioni d'uso, le licenze, i diritti di proprietà e l'utilizzo nel tempo delle risorse digitali.

Il Gruppo MAG ha elaborato uno schema di riferimento XML, basato sul modello METS (*Metadata Encoding and Transmission Standard*) utilizzato dalla *Library of Congress,* che predispone un set minimo di metadati gestionali al fine di una loro applicazione nei progetti di digitalizzazione. Lo schema è giunto alla versione 2.0.1[27] (marzo 2006), per l'implementazione dei metadati non solo descrittivi, ma anche di quelli amministrativi-gestionali nei progetti di digitalizzazione italiani; ha deciso, inoltre, un ulteriore approfondimento in relazione alla possibile adozione di terminologie alternative, quali ad esempio EAD (*Encoded Archival Description*) largamente impiegata negli archivi. Ad ogni modo, oggi lo schema MAG è lo standard utilizzato in Italia per le biblioteche digitali, ed è a questo modello che un progetto di digitalizzazione si deve riferire per l'applicazione di metadati ai beni digitalizzati.

---

[27] <http://www.iccu.sbn.it/genera.jsp?id=267>

Lo schema MAG è composto da varie sezioni e non è definito per particolari categorie di documenti ma per tipologie di oggetti digitali: immagini statiche e in movimento, documenti, suoni ecc. È quindi applicabile per qualsiasi oggetto da digitalizzare.

In sede di digitalizzazione, a ogni pagina digitale è associato un proprio set di metadati, indispensabile sia per la gestione quotidiana degli oggetti digitali (acquisizione, ricerca e gestione di archivi), sia per la loro conservazione a lungo termine.

Come già detto, lo schema MAG è basato sul modello logico-funzionale OAIS (*Open Archival Information System*), e ha l'obiettivo di dare le specifiche formali per la fase di raccolta e di trasferimento dei metadati e dei dati digitali nei rispettivi archivi; è inoltre indipendente da specifiche piattaforme hardware e software ed è espresso nella sintassi XML. Consente l'uso di metadati definiti e mantenuti da altre comunità (come *Dublin Core* e NISO), con l'aggiunta di altri metadati specifici. Di norma i MAG vanno associati alla risorsa digitale e comprendono informazioni generali sul progetto e sul metodo di digitalizzazione e altre informazioni relative al singolo oggetto digitale. Il set di elementi, più che per particolari tipologie di documento, è definito per tipologie di oggetti digitali (immagini statiche, testi prodotti con tipologia OCR, suono, audiovisivi ecc.). Lo schema MAG è composto da diverse sezioni, alcune generali e altre specifiche del contenuto digitale:

<gen>	informazioni generali sul progetto e sul tipo di digitalizzazione
<bib>	metadati descrittivi sull'oggetto analogico
<stru>	metadati strutturali
<img>	metadati specifici relativi alle immagini fisse
<ocr>	metadati specifici relativi al riconoscimento ottico del testo
<doc>	metadati specifici per oggetti digitali in formato testo che possono essere derivati o born digital
<audio>	metadati specifici per file audio
<video>	metadati specifici per file video
<dis>	metadati specifici per la distribuzione di oggetti bibliografici

Tabella 12.20: lo schema Mag

Secondo lo standard MAG, le sezioni obbligatorie, da un punto di vista tecnico, sono *<gen>* e *<bib>*. La sezione *<gen>* contiene le informazioni fondamentali relative all'istituzione responsabile del progetto di digitalizzazione, al progetto stesso, alla completezza o integrità del file, all'accessibilità dell'oggetto descritto. La sezione *<bib>* raccoglie i metadati descrittivi dell'oggetto.

Essendo un *application profile*, lo standard MAG interagisce e interopera con diversi standard internazionali di codifica di

metadati, con la possibilità di integrare elementi provenienti da più *namespace*. In alcuni casi, infatti, assume altri schemi di codifica (*Dublin Core* e NISO), in altri casi, invece, può essere trasformato in formati diversi (METS e MPEG-21).

12.2.2 Dublin Core

Dublin Core è un sistema di metadati costituito da un nucleo di elementi essenziali per la descrizione di qualsiasi materiale digitale. La specifica *Dublin Core* è diventata standard ISO15836 nel 2003 e rappresenta il sistema più diffuso a livello internazionale. Il comitato guida della Biblioteca Digitale Italiana, su indicazione del gruppo MAG, costituito presso l'ICCU e composto da personale tecnico di responsabili e tecnici di archivi, biblioteche e musei, ha raccomandato l'adozione di questo standard per l'implementazione dei meta dati descrittivi.

Il progetto della *Dublin Core Metadata Initiative* è nato e si è sviluppato nell'OCLC (*On line Computer Library Center*), la grande rete americana di servizi per le biblioteche. Nel 1995, durante una conferenza nella città di Dublin in Ohio, i partecipanti discusso sulla necessità di creare un insieme di strumenti condivisi per l'accesso alle risorse digitali. Lo scopo era quello di stabilire un nucleo (*core*) di elementi descrittivi che potessero essere forniti dall'autore o dall'editore dell'oggetto digitale, e inclusi in esso.

Lo standard *Dublin Core* definisce un insieme di elementi suddiviso in due livelli: *simple* e *qualified*. Il livello *simple Dublin Core* prevede quindici elementi fondamentali[28]:

1. **Titolo** (*title*), il nome dato alla risorsa o un nome con il quale la risorsa è formalmente conosciuta;
2. **Creatore** (*creator*), un'entità che ha la responsabilità principale della produzione del contenuto della risorsa: una persona, un'organizzazione o un servizio;
3. **Oggetto** (*subject*), l'argomento della risorsa. Può essere espresso da parole chiave o frasi chiave o codici di classificazione che descrivono l'argomento della risorsa;
4. **Descrizione** (*description*), una spiegazione del contenuto della risorsa: un riassunto analitico, un indice, un riferimento al contenuto di una rappresentazione grafica o un testo libero del contenuto;
5. **Editore** (*publisher*), Un'entità responsabile della produzione della risorsa, disponibile nella sua forma presente;
6. **Autore di contributo subordinato** (*contributor*), un'entità responsabile della produzione di un contributo al contenuto della risorsa;

[28] *Dublin Core Metadata Element Set* (versione 1.1), trad. it a cura dell'ICCU (Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche). <http://www.iccu.sbn.it/genera.jsp?id=116>.

7. **Data** (*date*), una data di 8 caratteri nella forma YYYY-MM-DD associata a un evento del ciclo di vita della risorsa. In genere, la data è associata alla creazione o alla disponibilità della risorsa;
8. **Tipo** (*type*), la natura o il genere del contenuto della risorsa: include termini che descrivono categorie generali, funzioni, generi o livelli di aggregazione per contenuto;
9. **Formato** (*format*), la manifestazione fisica o digitale della risorsa. Può includere il tipo di supporto o le dimensioni (grandezza e durata) della risorsa ed essere usato per determinare il software, l'hardware o altro apparato necessario alla visualizzazione o all'elaborazione della stessa;
10. **Identificatore della risorsa** (*identifier*), un riferimento univoco alla risorsa nell'ambito di un dato contesto, espresso per mezzo di una sequenza di caratteri alfabetici o numerici secondo un sistema di identificazione formalmente definito;
11. **Fonte** (*source*), un riferimento a una risorsa dalla quale è derivata la risorsa in oggetto;
12. **Lingua** (*language*), la lingua del contenuto della risorsa, espresso con un codice di lingua su due caratteri seguito opzionalmente da un codice di paese su due caratteri. Ad

esempio, *en* per l'inglese, *fr* per il francese, *it* per l'italiano o *en-uk* per l'inglese usato nel Regno Unito;

13. **Relazione** (*relation*), un riferimento alla risorsa correlata;
14. **Copertura** (*coverage*), l'estensione o scopo del contenuto della risorsa: la localizzazione spaziale (il nome di un luogo o le coordinate geografiche), il periodo temporale (l'indicazione di un periodo, una data o un range di date), una giurisdizione (ad esempio il nome di un'entità amministrativa);
15. **Gestione dei diritti** (*rights*), l'informazione sui diritti esercitati sulla risorsa. Comprende i diritti di proprietà intellettuale *Intellectual Property Rights* (IPR), copyright e vari diritti di proprietà. Se l'elemento *rights* è assente, non si può fare alcuna ipotesi sullo stato di questi o altri diritti in riferimento alla risorsa.

Il livello *qualified Dublin Core* aggiunge altri elementi e introduce un insieme di qualificatori che consentono di affinare la semantica degli elementi di base. I qualificatori consentono di descrivere con maggior dettaglio le informazioni fornite tramite gli elementi di base. Sono rappresentati dagli *element refinement*, elementi che specificano meglio determinate caratteristiche, e dagli *encoding*

scheme, insiemi di riferimento per l'interpretazione standard dei valori degli elementi.

Il livello *qualified Dublin Core* include:

1. **Pubblico** (*audience*), un insieme di entità per cui la risorsa risulta utile;
2. **Origine** (*provenance*), le informazioni su eventuali modifiche dalla creazione della risorsa;
3. **Utilizzo dei diritti** (*rightsholder*), i detentori i diritti di utilizzo della risorsa;
4. **Fini educativi** (*instructionalmethod*), il processo di apprendimento per il quale la risorsa è indicata;
5. **Modalità d'incremento** (*accrualmethod*), il metodo con cui la risorsa è aggiunta ad un insieme;
6. **Intervallo d'incremento** (*accrualperiodicity*), la periodicità di aggiunta di una risorsa ad un insieme;
7. **Politica d'incremento** (*acrualpolicy*), la politica che regola l'aggiunta della risorsa ad un insieme[29].

I vantaggi dell'adozione dello standard *Dublin Core* sono numerosi: è semplice e nasce per i catalogatori di risorse web; stabilisce una comune rete di dati concordanti nel loro significati e

[29] *Dublin Core Metadata Element Set*, op. cit. Mia la traduzione

valore e garantisce l'interoperabilità semantica; offre uno strumento utile per creare infrastrutture a livello internazionale ed è molto flessibile: permette di integrare e sviluppare la struttura dei dati con significati semantici diversi e applicati al contesto della specifica applicazione.

12.2.3 NISO e MIX

Il NISO *Technical Metadata for Digital Still Images Standards Committee* ha sviluppato delle linee guida per la creazione di metadati amministrativi e gestionali relativi alle immagini statiche, contenute nel *Data Dictionary - Technical Metadata for Digital Still Images*, il cui obiettivo è così espresso:

> The purpose of this data dictionary is to define a standard set of metadata elements for digital images. Standardizing the information allows users to develop, exchange, and interpret digital image files. The dictionary has been designed to facilitate interoperability between systems, services, and software as well as to support the long-term management of and continuing access to digital image collections[30].

A differenza del *Dublin Core*, lo standard NISO non comprende uno schema di codifica XML, ma si limita a fornire una serie di linee guida per la creazione di linguaggi di *markup*. Lo

---

[30] *Data Dictionary - Technical Metadata for Digital Still Images*, <http://www.niso.org/pdfs/DataDict.pdf>, p. 1.

standard non è ancora stato rilasciato formalmente e viene distribuito in *trial use*.

Lo schema MAG assume lo standard NISO, sviluppando un linguaggio di *markup* definito *NISO-MAG*, incluso al proprio interno tramite un *namespace* il cui prefisso distintivo è "niso.". È stata stabilita da parte del Comitato MAG la scelta di attendere una versione stabile dello standard prima di procedere all'aggiornamento dello schema di codifica.

La *Library of Congress*, insieme al *MARC Standards Office*, ha elaborato uno schema di codifica basato sulle linee guida del *Data Dicitionary NISO*, che è confluito nello schema di codifica NISO MIX, attualmente giunto alla versione 2.0[31].

12.2.4 METS

Lo schema METS, sviluppato dalla *Library of Congess*, è uno standard per codificare metadati descrittivi, amministrativi e strutturali relativi a oggetti all'interno di una biblioteca digitale, con le stesse finalità di MAG, ma che si propone più come "schema contenitore" che come risposta o punto di riferimento per la registrazione di metadati. METS può integrare al suo interno diversi schemi di codifica non predeterminati, mentre offre poche linee guida sulla modalità e gli oggetti da codificare: per le immagini statiche, METS consiglia l'adozione di NISO MIX; per la codifica

[31] <http://www.loc.gov/standards/mix>

di audio e video non fornisce alcuna indicazione.

METS è sicuramente uno schema con molte potenzialità e che ha incontrato l'interesse di numerose istituzioni in tutto il mondo. Il Comitato MAG ha elaborato un applicativo (basato su un foglio di stile XSLT) in grado di trasformare un file MAG in un file METS, in modo da coordinare progetti che adottano i diversi schema di codifica: il sistema di conversione portà garantire l'interscambiabilità e la condivisione dei dati a livello internazionale. In ogni caso MAG può essere usato come estensione di METS.

12.2.5 MPEG-21

L'ultimo nato fra gli schemi di codifica dei matadati è l'MPEG-21:

> MPEG-21 aims at defining a normative open framework for multimedia delivery and consumption for use by all the players in the delivery and consumption chain. This open framework will provide content creators, producers, distributors and service providers with equal opportunities in the MPEG-21 enabled open market[32].

Il cuore dell'MPEG-21 è il concetto di *digital item*. I *digital items* sono oggetti digitali strutturati che includono una

---

[32] <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>

rappresentazione standard, un'identificativo e dei metadati. Più concretamente, un *digital item* è costituito di una combinazione di risorse (come uno *stream* video, tracce audio, immagini statiche), metadati e strutture che descrivono le relazioni che intercorrono fra le risorse.

MPEG-21 si occupa della definizione dei requisiti tecnici fondamentali degli oggetti digitali e della possibilità di interazione da parte dell'utente con gli stessi. Il secondo di questi aspetti investe direttamente i metadati, grazie ai quali è infatti possibile interagire con l'oggetto digitale. A questo proposito, la parte seconda dello standard contiene le specifiche di uno schema di codifica per dichiarare la struttura e le caratteristiche dei *digital items*, denominato DIDL (*Digital Item Declaration Language*), le cui maggiori caratteristiche sono la flessibilità e l'interoperabilità. È in fase di studio la possibilità di realizzare un applicativo analogo a quello realizzato per METS per convertire i file MAG in file DIDL oltre a delle linee guida che consentano l'interoperabilità dei formati.

12.3 XML

XML, acronimo per *eXtensible Markup Language*, è un linguaggio di *markup* sviluppato dal *World Wide Web Consortium* W3C nel 1999, con lo scopo di creare uno standard per le applicazioni

web e per la gestione dei dati. Il linguaggio XML ha trovato un largo ambito di applicazione, dal settore dell'editoria digitale alla conservazione e l'interscambio dei dati.

Un linguaggio di *markup* è costituito da un insieme di marche o etichette (dette in gergo *tag*) che servono ad annotare e rappresentare una proprietà o la natura di un determinato testo o di un dato. XML consente di creare marche personalizzate che possono descrivere in maniera immediata ogni caratteristica del dato considerato. Inoltre, essendo garantito e sviluppato dal W3C, XML fornisce buone garanzie circa la sua durata e la possibilità di conservare i dati nel lungo periodo.

L'operazione di annotazione dei dati tramite le marche XML si definisce normalmente *marcatura*. Grazie alla possibilità di creare linguaggi personalizzati, molte istituzioni hanno creato i propri schemi di marcatura che hanno poi proposto alla comunità internazionale, lasciando aperta la possibilità di adottare uno di questi schemi e codificare i propri dati seguendone le direttive e le specifiche.

Pur essendo altamente flessibile e personalizzabile, il linguaggio XML stabilisce tre livelli di correttezza, dei quali è obbligatorio soltanto il primo. Un documento codificato in XML dovrebbe essere:

1. **ben formato**, rispettando le specifiche della sintassi XML. Questo aspetto è verificato automaticamente dal computer grazie ad analizzatori sintattici, detti *parser*;

2. **valido**, rispettando i vincoli dello schema di codifica prescelto. Questo aspetto è verificato dai *parser* validanti;

3. **semanticamente coerente**, quando il documento rispetta la semantica dello schema di codifica prescelto.

12.3.1 Sintassi XML

Elementi

L'unità base di XML è l'*elemento*: un'unità di significato considerata in quanto componente strutturale di un documento. Un documento XML può essere descritto come una sequenza organizzata di elementi, tutti racchiusi da un elemento radice (o elemento *root*). Diversi tipi di elementi assumono diversi nomi.

Il nome di un elemento è detto *generic identifier* o GI. Per marcare un elemento è necessario introdurre una particolare sequenza di caratteri per segnare l'inizio dell'elemento (detto *start tag*), formato dal nome dell'elemento racchiuso fra parentesi angolate, e un'altra sequenza di caratteri per segnarne la fine (detto

end tag), formato dal nome dell'elemento preceduto da uno slash e racchiuso fra parentesi angolate. Un semplice esempio:

```
<elemento>Questo è un elemento</elemento>
```

in questo caso la codifica <*elemento*> indica l'inizio dell'elemento, e la codifica </*elemento*> ne indica la fine.

Un elemento può anche essere vuoto, il che significa che potrebbe non avere alcun contenuto. La sintassi di un elemento vuoto può assumere due forme:

1. uno *start tag* seguito da un *end tag*, non inframmezzati da nessun carattere, nemmeno da uno spazio:

```
<elementoVuoto></elementoVuoto>
```

2. il nome dell'elemento seguito da uno slash, racchiuso da parentesi angolate:

```
<elementoVuoto />
```

Un documento XML può consistere in una serie di elementi l'uno dopo l'altro racchiusi dall'elemento radice. Ad esempio:

```
<radice>
    <titolo> Questo è un elemento titolo</titolo>
    <p>Questo è un paragrafo</p>
    <nota>Questa è una nota</nota>
</radice>
```

Oppure, nel caso più frequente, gli elementi possono essere annidati gli uni dentro gli altri, come:

```
<radice>
    <capitolo>Questo è un capitolo,
        <titolo> Questo è un elemento titolo</titolo>
        <p>Questo è un paragrafo</p>
    </capitolo>
</radice>
```

Le relazioni tra gli elementi in XML sono essenzialmente relazioni gerarchiche e ordinali. XML si basa su una struttura gerarchica in cui non esistono sovrapposizioni tra elementi. Questo significa che la struttura astratta di un documento XML deve essere rappresentabile mediante un grafo, nel quale a ciascun nodo corrisponde un elemento e a ogni ramo verso il basso una relazione di inclusione. Tutti gli elementi devono essere racchiusi all'interno dell'elemento *root* o all'interno di elementi a loro volta racchiusi nell'elemento root; gli elementi non possono intrecciarsi (*overlapping*) gli uni agli altri. Vale a dire che se l'elemento B è viene aperto all'interno dell'elemento A, B deve essere chiuso prima che si chiuda A.

In una citazione bibliografica, ad esempio:

```
<citazione>
    <autore>Federico De Roberto</autore>
    <titolo>I Viceré</titolo>
    <editore>Einaudi</editore>
    <data>2006</data>
    <soggetto/>
</citazione>
```

Per definire i rapporti reciproci fra gli elementi, si usa normalmente la metafora dei rapporti di parentela. Quando un elemento è contenuto direttamente dentro un altro, si dice che quell'elemento è *figlio* dell'altro. Nell'esempio precedente, l'elemento autore è *figlio* dell'elemento citazione; analogamente, citazione è *genitore* di autore. Se un elemento è figlio di un figlio di un altro elemento, si dice che ne è il *discendente*, mentre l'elemento da cui discende è detto *antenato*.

Attributi

Gli attributi servono a specificare meglio le caratteristiche degli elementi e quindi dei dati. Tramite gli attributi è possibile descrivere un particolare stato di un elemento o dare maggiori informazioni sullo stesso. Gli attributi sono costituiti da un nome e da un valore legato al nome da un segno di uguale, racchiuso fra virgolette secondo la sintassi:

```
nome="valore"
```

L'attributo deve essere dichiarato all'interno dello *start tag* o in un elemento vuoto, con il valore racchiuso fra virgolette semplici o doppie.

Possiamo dire, ad esempio, se una particolare citazione riguarda una monografia oppure il contributo di una miscellanea:

```
<citazione tipo="monografia"> >
```

```
    <autore>Federico De Roberto</autore>
    <titolo>I Viceré</titolo>
    <editore>Einaudi</editore>
    <data>2006</data>
    <soggetto/>
</citazione>
```

Commenti

A volte può essere utile utilizzare commenti nei file XML, note che vengono inserite dal codificatore allo scopo di annotare una caratteristica non strutturale o per ricordare qualcosa. I commenti sono ignorati dai *parser* XML e possono essere inseriti in qualunque punto del documento, purchè non siano posti all'interno dei tag.

I commenti si aprono con la stringa "<!—" e si concludono con "-->"

```
<citazione tipo="monografia"> >
    <autore>Federico De Roberto</autore>
    <titolo>i Viceré</titolo>
    <editore>Einaudi</editore>
    <data>2006</data>
    <!-- attenzione: ricordarsi di inserire il soggetto -->
    <soggetto/>
</citazione>
```

Namespaces

Per utilizzare all'interno di un documento XML alcuni elementi che hanno una particolare provenienza, definiti in un qualche schema di codifica, XML adopera i *namespace*. Un *namespace*

è costituito da un prefisso distintivo da dichiarare all'interno di un qualsiasi elemento di un documento XML, usando un attributo standard disponibile per tutti gli elementi XML. L'attributo si chiama *xmlns* (abbreviazione di *XML NameSpace*) e deve essere composto dal prefisso distintivo del *namespace* da adottare, seguito da simbolo di due punti.

La dichiarazione di un *namespace* si effettua per mezzo di un attributo standard speciale previsto dalle specifiche XML, “xmlns:prefisso”. Al valore è assegnato dall’indirizzo URL della documentazione dello schema di riferimento.

Per utilizzare, ad esempio, gli elementi definiti dallo schema *Dublin Core* per codificare una citazione bibliografica sarà sufficiente usare il namespace “dc”:

```
<citazione xmlns:dc="http://purl.org/dc/elements/1.1/">
   <dc:creator>Federico De Roberto</dc:creator>
   <dc:title>I Viceré</dc:title>
   <dc:publisher>Einaudi</dc:publisher>
   <dc:date>2006</dc:date>
   <dc:subject/>
</citazione>
```

12.3.2 Gli schema W3C

Dichiarare uno schema di codifica per un documento XML significa formulare una serie di regole che definiscano gli elementi da usare e il modo in cui tali elementi possano relazionarsi: la

possibilità di essere contenuti da altri elementi, il numero di volte in cui possono essere usati, quali attributi dovranno essere utilizzati potranno essere e se dovranno o meno avere un certo contenuto.

Il metodo più utilizzato per dichiarare uno schema di codifica è stato attraverso una *document type definition* (o DTD), una sintassi ereditata direttamente dal linguaggio SGML, da cui deriva l'XML. La DTD è costituita da una sintassi estremamente semplice, ma dà luogo a considerevoli limitazioni, soprattutto per quanto riguarda la definizione delle diverse tipologie di elementi e delle occorrenze degli stessi.

Alla *document type definition* si sono attualmente aggiunti gli *schema*, secondo una raccomandazione del *World Wide Web Consortium* proposta sin dalla nascita del linguaggio XML, nel 1999. Questi *schema*, attraverso una sintassi diversificata, raggiungono più o meno gli stessi obiettivi nel superamento dei limiti posti dalle DTD.

Lo schema di codifica MAG addotta le specifiche W3C.

13. Conservazione

In accordo con la definizione l'Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche, l'obiettivo della conservazione dei documenti intesa in senso tradizionale è

> assicurare la stabilità nel tempo dei supporti fisici, in modo da assicurare l'accesso all'informazione da essi veicolata[33].

In primo piano è dunque posta l'attenzione per l'accesibilità in tempi futuri, un argomento che l'avvento del digitale porta più che mai in primo piano. I nuovi fattori di rischio per la conservazione del patrimonio culturale in formato elettronico coinvolgono fattori quali l'obsolescenza dei formati elettronici, la transitorietà dei supporti fisici di archiviazione e la caducità dei dispositivi hardware. Se l'innovazione tecnologica permette di raggiungere risultati sempre più avanzati, di fatto mette a rischio l'accesibilità in tempi successivi se non si fa ricorso a specifici accorgimenti in tutti gli ambiti d'interesse: sono coinvolti in questo processo di "invecchiamento precoce" i documenti elettronici,

---

[33] Istituto Centrale per il Catalogo Unico, *Conservazione dei documenti digitali*, Iccu, <http://www.iccu.sbn.it/genera.jsp?s=75&l=it>

quelli ad accesso locale e remoto, i *born digital* e quelli derivati per digitalizzazione da precedenti documenti cartacei.

Un progetto di digitalizzazione non si conclude, dunque, con l'acquisizione dei documenti e le attività connesse alla diffusione: senza una scrupolosa cura nei confronti della conservazione dei beni digitali nel lungo periodo si rischia di vanificare tutto il lungo e impegnativo processo di lavoro.

Anche in questo caso, quindi, sono stati istituiti gruppi istituzionali qualificati per studiare e proporre soluzioni al problema. Il caso è stato discusso inizialmente all'interno di due importanti conferenze: la Risoluzione del Consiglio d'Europa con il tema *Preservare la memoria del futuro, preservare i contenuti digitali per le generazioni future* del 25 giugno 2002 e la Conferenza internazionale di Firenze del 16 e 17 ottobre 2003. Con la collaborazione della Commissione europea, della Presidenza italiana e dei progetti ERPANET (*Electronic Resource Preservation and Access Network*) e Minerva, è stato istituito un *Gruppo di lavoro per la conservazione delle memorie digitali* che ha stabilito tre obiettivi da raggiungere in ambito europeo:

1. delineare il quadro delle iniziative in corso e favorire lo scambio di "buone pratiche";
2. definire una agenda di priorità come punto di partenza per un piano d'azione condiviso dagli Stati-membri;
3. definire le basi per la formazione di una rete europea e per lo

sviluppo di iniziative nazionali[34].

In una prima fase di lavoro, il gruppo si è impegnato nel favorire un processo di cooperazione tra gli Stati europei, stabilendo la coordinazione per un processo di efficacia e incremento delle attività di conservazione digitale. I risultati sono stati presentati alla Conferenza internazionale di Firenze del 2003 con la denominazione di *Firenze Agenda*, firmata da 27 paesi.

Il passo successivo per la promozione delle pratiche di conservazione è stata la nascita nell'aprile del 2006 del progetto europeo DPE (*Digital Preservertion Europe)*, finanziato dalla Commissione europea, con la partecipazione della Direzione Generale per i beni librari, gli istituti culturali e il diritto d'autore del Ministero per i Beni e le Attività Culturali e il contributo scientifico dell'Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane. Il progetto è attualmente in corso.

13.1 Storage dei dati

Il processo di digitalizzazione produce una quantità di dati digitali che varia sensibilmente in base al numero di documenti da scansire e dal numero di postazioni di lavoro operative. Un sistema di digitalizzazione che lavori in parallelo su più supporti può

[34] Ivi

arrivare a generare quotidianamente decine di gigabyte. Garantire la conservazione e l'accesso alle informazioni digitali significa garantire l'affidabilità dei supporti di archiviazione, delle architetture hardware e software e delle infrastrutture che permettono il trasferimento dei dati. È essenziale stabilire e misurare le prestazioni e la sicurezza del sistema di archiviazione per prevenire l'insorgere di guasti e quindi la perdita dei dati.

13.1.1 Memorie offline

Le soluzioni tradizionali basate su sistemi di memorizzazione locali sono molteplici. Le architetture più adatte a conservare i dati acquisiti da un progetto di digitalizzazione sono:

Cd-rom. Sono supporti affidabili, anche in termini di permanenza dei dati. Il processo di acquisizione, masterizzazione, ed eventuale duplicazione è economico e molto semplice. Sono limiti oggettivi la scarsa quantità di dati archiviabili su singolo supporto e, di conseguenza, la necessità di organizzare e acquisire un sistema, automatico o manuale, di gestione degli stessi.

Dvd-rom. Possiede una maggiore capacità rispetto al CD. Le dimensioni dei DVD di produzione industriale sono di quattro tipi: lato unico e singolo strato (4.7 gb), lato unico e doppio strato (8.5

gb), due lati e singolo strato (9.4 gb) e due lati e doppio strato (17 gb).

Dmd. si tratta di una tecnologia di nuova generazione che si annuncia molto promettente. Il *digital multilayer disk* ha le dimensioni di un comune cd, ma le informazioni sono archiviate in modo stratificato, così da ottenere una capacità complessiva di 140 gigabytes (ma si parla anche di un prossimo sviluppo fino alla capacità di un terabyte, che corrisponde al contenuto di ben 1.500 CD). La velocità di trasferimento è elevatissima: un gigabyte al secondo. Il prodotto viene annunciato di prossima commercializzazione.

Hard disk esterni. Permettono l'immagazzinamento di una notevole quantità di dati, fino a qualche terabyte. Offrono il vantaggio di soddisfare esigenze di mobilità, permettendo il salvataggio dei documenti da diverse postazioni su un unico supporto. Adottano tecnologie di connessione largamente condivise, come porte USB, FireWire, SCSI, eSATA

13.1.2 Immagazzinamento ridondante

È necessario garantire la produzione di più copie di ogni documento digitalizzato, da archiviare preferibilmente su supporti conservati in luoghi diversi. L'immagazzinamento ridondante

consiste proprio nella creazione di più cloni dei dati digitali ottenuti. È consigliabile archiviare le copie su supporti fisici differenti, anche di diverso formato, al fine di garantire la reperibilità di almeno una copia di sicurezza in caso di guasti dei dispositivi di archiviazione.

La clonazione ridondante delle informazioni digitali dovrebbe essere messa in pratica contemporaneamente, o immediatamente dopo, l'acquisizione dei documenti in formato digitale. È possibile mettere in atto l'immagazinamento ridondante manualmente, attraverso il lavoro di un operatore che si occupa di clonare i dati sui vari supporti; attraverso un raid, un sistema che permette di combinare un insieme di dischi in una sola unità logica, gestendo autonomamente la copia dei dati sui vari drive; via software, attraverso sistemi di *backup* automatizzati che consentono la gestione gerarchica dei volumi.

13.2 Digital preservation

La conservazione del patrimonio digitale nel tempo è un argomento attualmente molto discusso. La velocità di evoluzione delle tecnologie informatiche, che coinvolge in modo uniforme formati, supporti di immagazzinamento dei dati, software applicativi e sistemi operativi, ha come conseguenza un grado di obsolescenza dei prodotti elettronici molto più elevato rispetto a quelli analogici e tempi più brevi di vita dei formati elettronici.

Non esistono ancora dati certi sulla durata dei differenti supporti fisici di memorizzazione dei dati, come hard disk, nastri magnetici, CD, CD-R e RW e DVD-R e *double layer*. Il tempo di vita per i CD stimato dalle case produttrici è di alcune decine di anni. Altri organismi internazionali, come la *Commission on preservation and access* negli USA e la Biblioteca Nazionale d'Australia, ipotizzano invece una durata attestata tra i cinque e gli otto anni.

Il progetto canadese InterPARES (*International Research on Permanent Authentic Records in Electronic Systems*)[35], portato avanti dalla *University of British Columbia*, ha rilevato però che può essere conservata la capacità di recuperare l'informazione contenuta da oggetto digitale, tramite opportune strategie di migrazione ed emulazione. La *Carta per la conservazione del patrimonio digitale* dell'UNESCO[36] ne individua quattro tipologie:

1. **Refreshing**. È una tecnica volta a copiare periodicamente i dati su nuovi supporti per mantenere i bit in buono stato. Tuttavia, questa operazione non salvaguarda dall'obsolescenza dei codici nativi dell'oggetto digitale. Si consiglia un refreshing dei dati archiviati su CD o DVD ogni cinque/otto anni;

---

[35] <http://www.interpares.org>

[36] <http://www.iccu.sbn.it/PDF/carta_UNESCO_eng.pdf>

2. **Migrazione e conversione dei dati**. Consiste nel trasferimento sistematico dei dati verso una configurazione hardware e software più aggiornata, al fine di contrastare l'obsolescenza dei formati nativi dei file. È consigliata una ricodifica periodica ogni cinque anni affinché i dati possano essere letti dalle nuove generazioni di hardware e software;

3. **Conservazione dell'ambiente originario di produzione.** È consigliabile conservare le configurazioni hardware e gli ambienti software (sistemi operativi e pacchetti applicativi) utilizzati per produrre i dati digitali per permetterne l'accesso in tempi successivi. Sarebbe opportuno conservare un esemplare funzionante per ogni generazione di macchine utilizzata per la produzione dei dati. Questa operazione, onerosa in temi di dispendio economico e di spazi, può essere affidata a istituzioni centralizzate a livello regionale o di sistema.

13.3 Creazione delle condizioni per l'emulazione.

Si basa sulla conservazione insieme al documento digitale delle informazioni relative alle configurazioni hardware e software che lo hanno prodotto, in modo da poter emulare su un computer di generazione successiva un ambiente analogo a quello in cui il

dato è stato generato. Ciò si può ottenere installando su un computer tecnologicamente aggiornato le piattaforme native o utilizzando appositi programmi in grado di emulare sistemi operativi e software obsoleti.

I supporti digitali, così come quelli analogici, sono soggetti a fattori di deterioramento fisico. Elementi ambientali, quali calore, luce, polvere, microrganismi, agenti chimici possono comprometterne il funzionamento. In linea di massima, le condizioni ambientali ideali per la conservazione dei supporti digitali sono una temperatura compresa tra i 5 e i 20 gradi centigradi, un tasso di umidità relativa compreso tra il 30% e il 50%, un gradiente massimo di temperatura di 4° l'ora e un gradiente massimo d'umidità relativa del 10% l'ora.

14. La digitalizzazione del *Voyage pittoresque des iles de Sicile, de Malte et de Lipari* di Jean Houel

14.1 Digitalizzare un libro antico

Nel caso dei libri antichi, il processo di digitalizzazione ha primo il compito di concentrarsi versante della conservazione: sia per conservare una copia digitale che "fissi" nel tempo le condizioni di testi che, inevitabilmente, sono destinati – anche se nel lungo periodo e con tutte le precauzioni del caso – al decadimento fisico, al deteriotamento della carta e del supporto. L'accesso alla copia digitale di un libro antico garantisce inoltre una possibilità estesa di lettura, prima impossibile considerate le restrizioni legittime alla consultazione per cercare di garantire la migliore conservazione nel tempo. I libri antichi sono chiaramente gli oggetti più rari tra quelli custoditi nelle biblioteche e negli archivi, i meno disponibili e, probabilmente, anche i più preziosi; diventano così, nelle migliori ipotesi, relegati alla consultazione per studi molto specifici o di importanza riconosciuta.

14.2 Il libro antico in formato elettronico: il formato immagine, il valore iconico

Così come per i manoscritti e tutti gli altri documenti rari, nei libri antichi gli elementi di testo decorativo e figurativo, insieme alle caratteristiche del supporto fisico possono essere a loro volta importanti nello studio del testo. Il tipo di carta, la legatura, gli elementi di *editing*, i set di caratteri utilizzati sono spesso a loro volta oggetto di studio; in ogni caso, ogni pagina di un libro antico è un oggetto da digitalizzare con la massima cura per far sì che si produca una copia quanto più possibile fedele all'originale. L'attenzione non va posta soltanto al testo, ma all'immagine nel suo complesso, dove l'immagine rappresenta la riproduzione esatta dell'impaginazione, degli elementi grafici, insieme al testo.

Nel caso, poi, di libri antichi illustrati o miniati l'argomento si fa ancora più interessante e utile, perché offrono numerose possibilità alternative di fruizione rispetto al modello del libro tradizionale. Come spiega Angela Nuovo a proposito della gestione biblioteconomica e del libro antico in rete:

> I progetti di digitalizzazione più affascinanti sono quelli che mescolano vari materiali e che dimostrano, ad esempio, quante immagini siano conservate nelle biblioteche, immagini che la gestione biblioteconomica tradizionale, affinatasi in secoli di descrizione bibliografica del libro come supporto di un testo, non è in grado di valorizzare o di indicizzare adeguatamente. Il formato

digitale è l'ideale per organizzare materiali diversi intorno ad alcune idee ordinatrici[37].

14.3 Il *Voyage pittoresque des iles de Sicile, de Malte et de Lipari* di Jean Houel

Il *Voyage pittoresque des iles de Sicile, de Malte et de Lipari* è un testo di Jean Houel pubblicato tra il 1782 e il 1787. Una copia originale della prima edizione del volume è custodita all'interno della sezione di Archeologia della Biblioteca della Facoltà di Lettere e Filosofia, sita in via Biblioteca 4. A causa del valore e dello stato di conservazione dei 4 tomi, il libro non è disposto alla consultazione.

14.4 Il progetto di digitalizzazione

Insieme l'interesse per alcune tavole manifestato dal prof. Giuseppe Giarrizzo per l'inserimento nel volume *Catania. La grande Catania, la nobiltà virtuosa, la borghesia operosa*[38], pubblicato per Sanfilippo Editore, ha costituito lo spunto per la proposta di digitalizzazione dell'intero volume. L'interessante racconto del viaggio nei luoghi siciliani, di Malta e delle isole Eolie, insieme al

[37] Angela Nuovo, *Libri antichi in rete* in *Soluzioni informatiche e telematiche per la filologia*, atti del convegno (Pavia, 30-31 marzo 2000), pubblicazione telematica a cura della Facoltà di Lettere e Filosofia dell'Università degli Studi di Pavia, disponibile all'indirizzo <http://lettere.unipv.it/dipslamm/pubtel/Atti2000/Nuovo0.htm>, consultato in data 26/10/2011.

[38] Enrico Iachello (a cura di), *Catania. La grande Catania, la nobiltà virtuosa, la borghesia operosa*, Domenico Sanfilippo Editore, Catania 2011.

nucleo di 264 disegni, offre un ampio sguardo sulla rappresentazione dei luoghi e della rappresentazione della Sicilia, di Malta, di Lipari da parte di uno dei fondamentali viaggiatori del *Grand Tour* settecentesco.

14.5 Il libro, le condizioni dei tomi

I quattro tomi che costituiscono l'opera sono volumi di grandi dimensioni, superiori al formato A3 per pagina, e hanno dunque richiesto di particolari accorgimenti per la digitalizzazione.

Le condizioni del libro antico, stampato su carta pregiata, ha permesso di mantenere nei suoi tre secoli di vita un buono stato di conservazione; tuttavia è stato necessario maneggiare i quattro tomi con estrema cura, poiché le legature risultavano talvolta compromesse e molto fragili. Le pagine illustrate sono state ottenute attraverso il processo settecentesco dell'acquatinta, tecnica di incisione indiretta complementare dell'acquaforte, ottenuta con una lastra granulata e granelli di bitume. Le veline, interposte tra le pagine illustrate, hanno consentito un'ottima conservazione dei pigmenti e della carta.

Figura 14.33: Una velina sulla pagina destra del I tomo del Voyage pittoresque des iles de Sicile, de Malte et de Lipari di Jean Houel

14.6 L'acquisizione delle immagini

Le dimensioni già citate dei quattro volumi, insieme alla delicatezza dei materiali, hanno richiesto una complessa apparecchiatura per la scansione ottica. Il comune scanner piano, infatti, avrebbe comportato un'azione difficilmente sopportabile da questi volumi; ognuno di questi, infatti, prima di ogni scansione sarebbe dovuto essere essere aperto sulle pagine da scansionare, posizionato a faccia in giù, pressato con forza sul piano di appoggio (per ridurre al minimo l'effetto curvatura della pagina e le ombre prodotte) e centrato opportunamente.

Queste operazioni manuali, ripetute per centinaia di volte su ogni volume, avrebbero provocato non solo un'enorme perdita di tempo, ma anche il prevedibile sfaldamento della rilegatura del volume. Per risolvere il problema è stato utilizzato uno scanner planetario, che permette di "rovesciare" il sistema di ripresa, eseguendolo dall'alto.

Lo scanner planetario, di proprietà della biblioteca della Facoltà di Lettere e Filosofia, ha permesso di ridurre al minimo lo "stress" della scansione applicato ai volumi, posizionando ciascun tomo una sola volta, a faccia in su, centrandolo rispetto al piano di ripresa. L'unico movimento a cui è stato sottoposto ciascun volume è stato quello, quasi impercettibile, di sfogliarne le pagine prima di avviare la scansione vera e propria, che consiste in una "fotografia" digitale ad altissima risoluzione (600 dpi). La ripresa dall'alto, condizione indispensabile per manipolare al minimo i volumi, insieme alla presenza di un piano basculante che ha consentito, una volta appoggiato il libro, di compensare il dislivello tra le due parti del volume in modo da metterle sempre in linea rispetto al piano di ripresa, e la correzione ottica dell'immagine da parte del software dello stesso scanner, necessaria a eliminare ombre, disallineamenti e curvature della pagina, hanno permesso di ottenere una copia digitale dei volumi di altissima qualità e perfettamente identica alla copia originale.

Figura 14.34: Lo scanner planetario utilizzato per il progetto di digitalizzazione

Figura 14.35: La fase di acquisizione delle immagini con lo scanner planetario di uno dei tomi

Gli interventi, ripetibili una tantum a causa del breve periodo di prestito del *Vojage,* sono stati quanto più accurati possibili: bilanciamento del bianco ripetuto *ex novo* per ogni pagina, archiviazione ridondante per evitare il pericolo di perdita dei dati digitali e, non ultimo, accurata conservazione del volume originale all'interno di un armadio blindato appositamente custodito.

Sono stati scrupolosamente rispettati standard e criteri di digitalizzazione conformi alle *Linee guida* e alle normative fissate dal Ministero. La modalità organizzativa e tecnica di acquisizione delle immagini è stata condotta con particolare riguardo alla qualità della scansione, ai formati per l'archiviazione, alle convenzioni per i nomi dei file e in generale all'alta qualità dei file digitali.

14.6.1 Il master

I file master sono stati acquisiti in formato TIFF non compresso a una risoluzione di 600 dpi. I requisiti di scansione adottati, la profondità di colore e i formati di output sono quelli individuati tramite l'adattamento della *Normativa sull'acquisizione digitale*, redatta dall'ICCD nel '98, agli attuali standard *de facto*.

A causa delle dimensioni notevoli e fuori standard e alla delicatezza del supporto fisico è stato scelto di procedere acquisendo delle immagini a doppia pagina per maneggiarlo il meno

possibile. I file master si presentano dunque con due pagine in ciascun file, memorizzati con una nomenclatura che prevede un codice univoco, composto da caratteri numerici e alfanumerici, che permette di riferirsi alla pagina o alla parte originale del manoscritto.

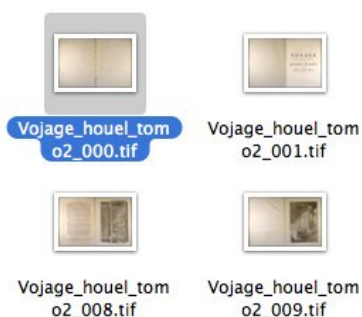



Figura 14..36: La nomenclatura di alcuni file tiff master del secondo tomo del Voyage pittoresque des iles de Sicile, de Malte et de Lipari di Jean Houel

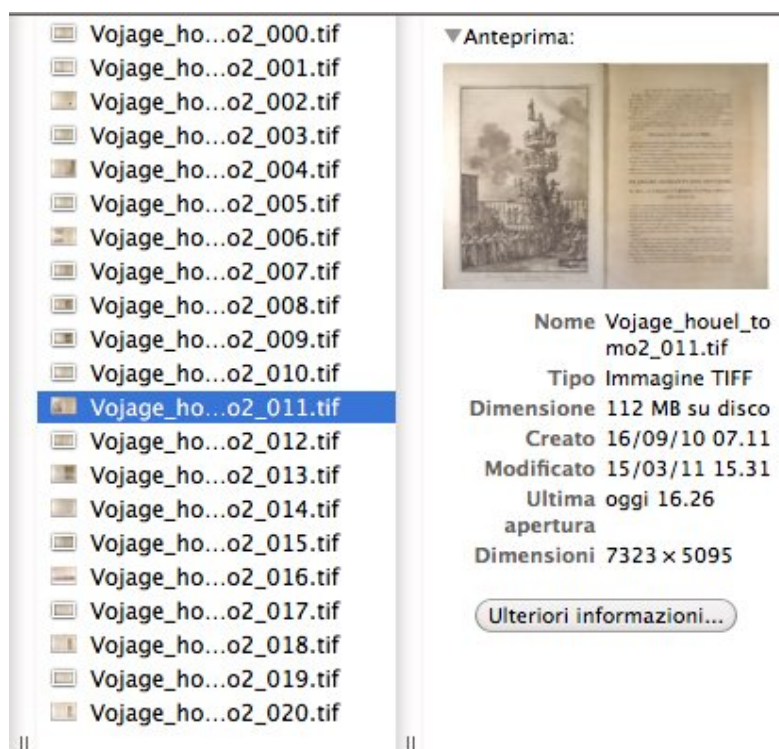



Figura 14.37: Schermata che riproduce i files tiff master del secondo tomo del Voyage pittoresque des iles de Sicile, de Malte et de Lipari di Jean Houel

La prima voce della nomenclatura si riferisce al titolo abbreviato dell'opera (*Vojage*), la seconda al nome dell'autore (*Houel*); segue il numero del tomo e infine un numero a tre cifre che serve a tenere in ordine progressivo le immagini digitalizzate. Non è stato possibile stabilire una corrispondenza tra numerazione progressiva nella digitalizzazione e numerazione di pagina dei volumi, sia a causa dell'inserimento a doppia pagina, sia a causa del fatto che sono presenti molte pagine vuote (bianche) non numerate, che costituiscono perlopiù il retro delle tavole illustrate. Le stesse illustrazioni sono ordinate con una numerazione romana progressiva all'interno del testo che non rispetta la numerazione principale di pagina.

Le immagini acquisite sono state anche raccolte in un file PDF diverso per ciascun tomo.

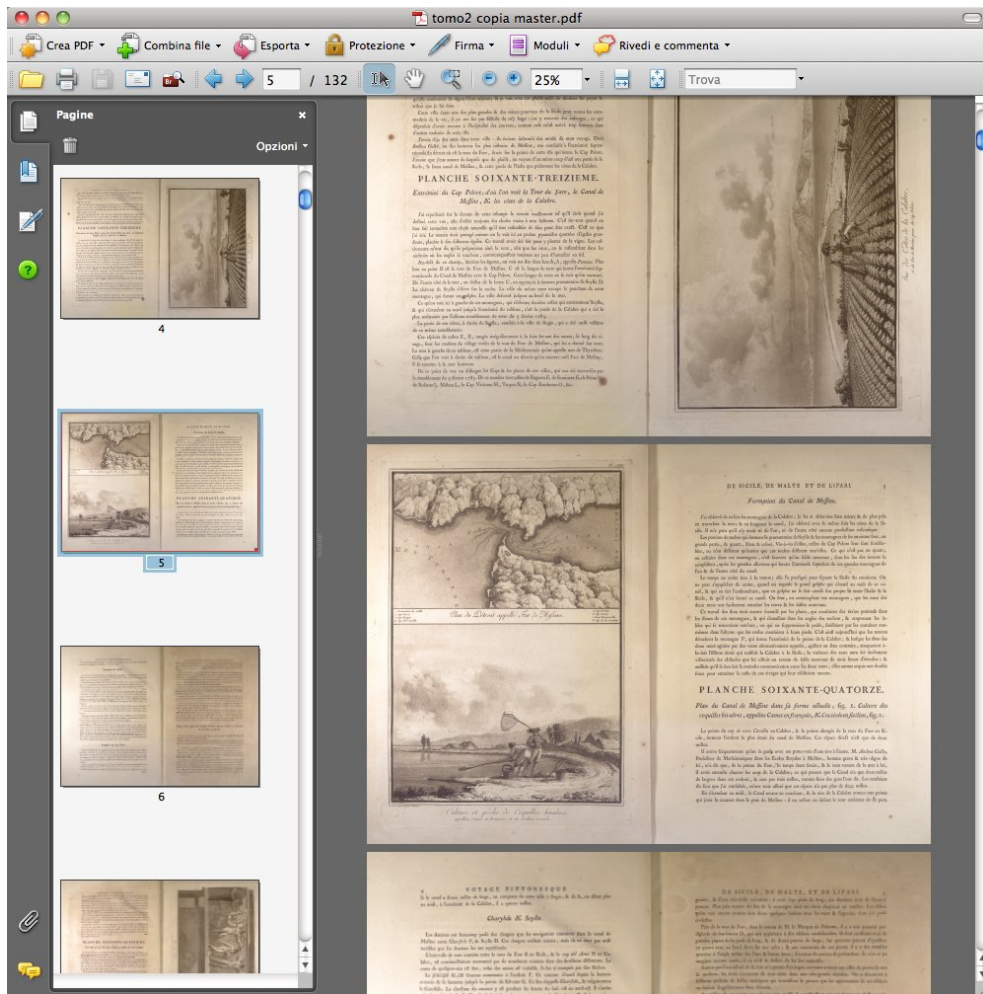



Figura 14.38: Il file pdf master del tomo II

14.7 Copie di consultazione

A completamento della copia master dell'intero processo di digitalizzazione di è proceduto con la creazione di una copia di consultazione. Le immagini acquisite ad altissima risoluzione sono

così state ricampionate da 600 dpi a 300 dpi e sono state lavorate in modo tale da separare le singole pagine, rispetto alla doppia pagina intera della copia master ed è stato adottato il formato JPG a qualità massima. Anche in questo caso, oltre ai singoli file immagine, è stato creato un file PDF per ogni singolo tomo.

Figura 14. 39: Il file pdf della copia di consultazione del secondo tomo.

14.8 Le tavole

Le acquetinte del *Voyage pittoresque des iles de Sicile, de Malte et de Lipari* costituiscono un nucleo di contenuto fondamentale all'interno dei quattro tomi. Sono il corredo d'immagine fondamentale per la testimonianza scritta del viaggio e la documentazione dei luoghi visitati e raccontati, ma sono ben lontane dall'avere valore esclusivamente come visione artistica o documentaria del territorio.

Figura 14.40: La Tavola I del I tomo

Figura 14.41: La Tavola 5, Veduta interna del Tempio di Segesta

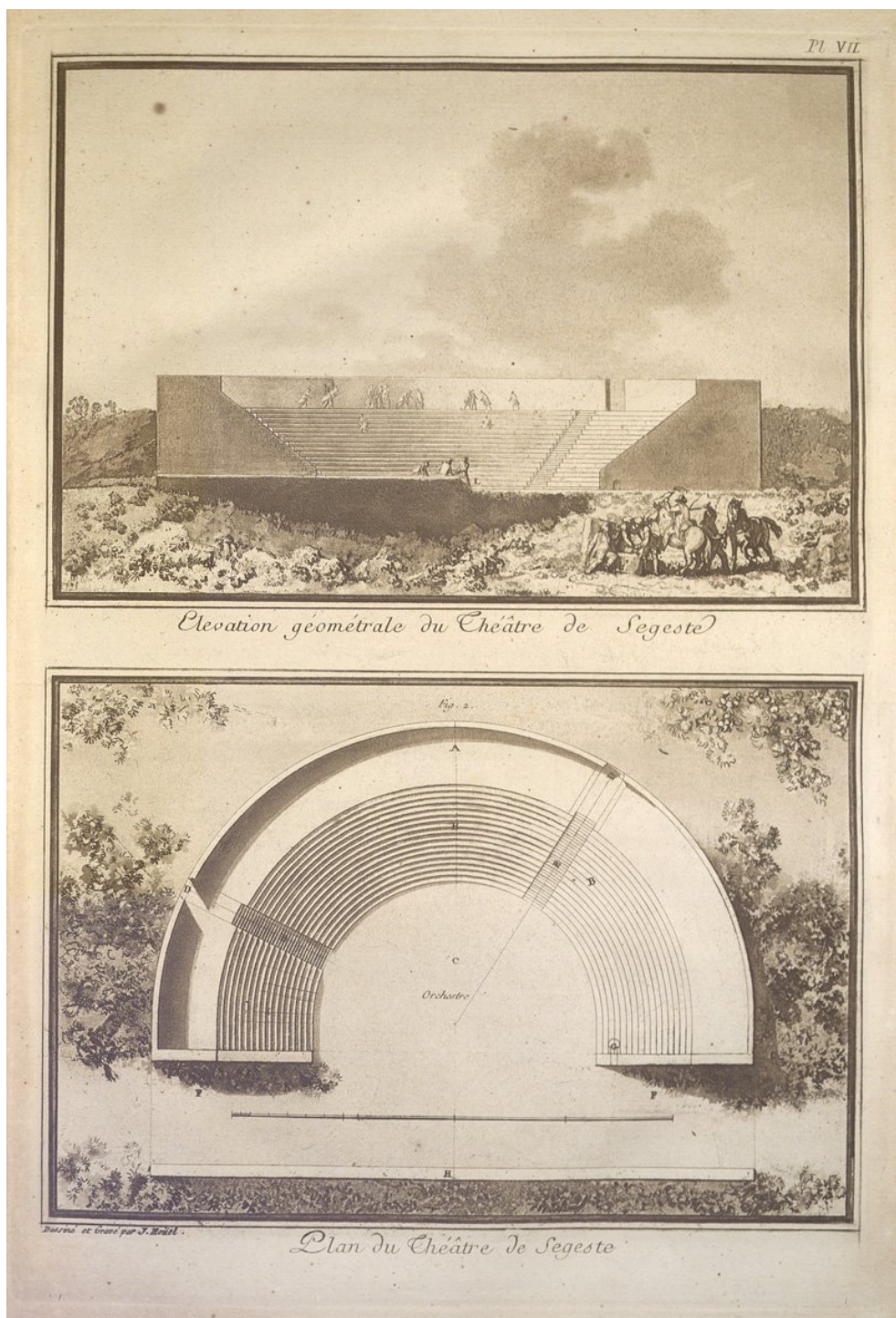



Figura 14.42: Tavola VII: rilievi architettonici del tempio di Segesta

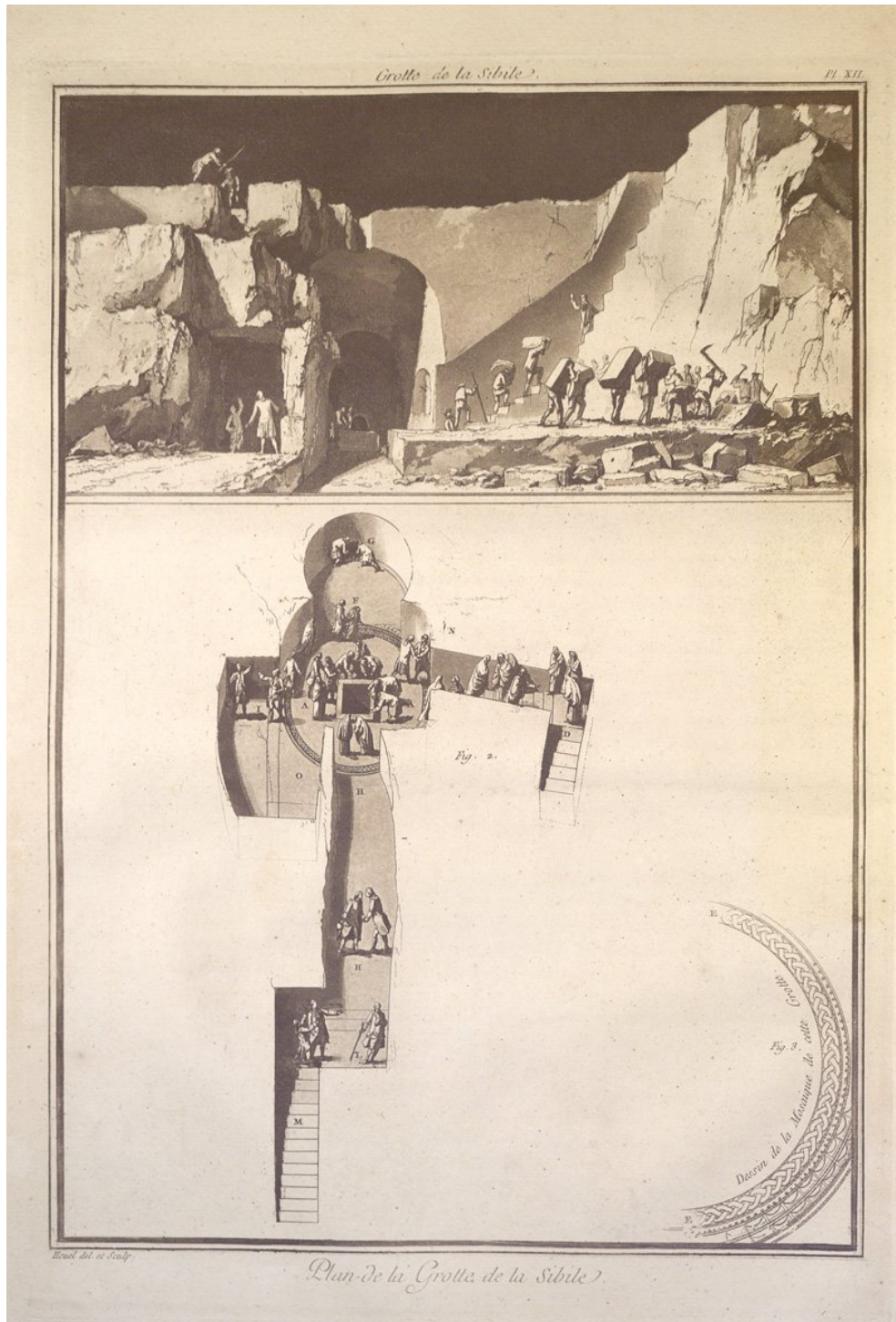



Figura 14.43: Tavola XII, rilievi della grotta della Sibilla

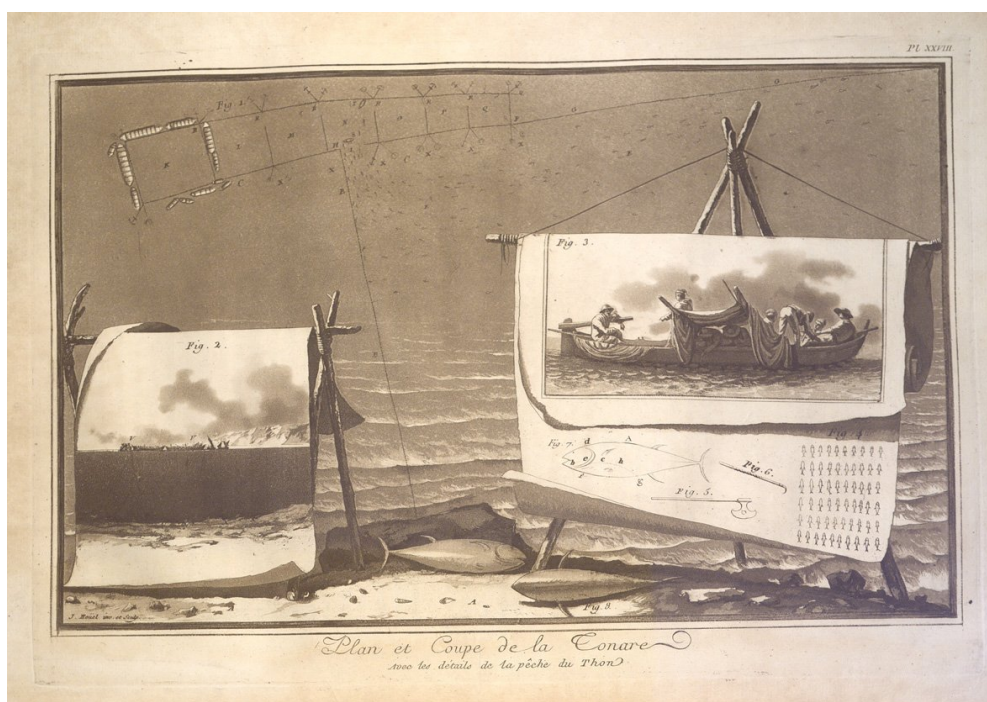



Figura 14.44: Tavola XXVIII, dettagli sulla pesca del tonno e sulle tonnare

Le tavole di Jean Houel sono il frutto dell'interpretazione da parte del viaggiatore francese di un sistema di relazioni tra territorio e società. Rappresentano i luoghi come "fenomeni" descritti attraverso un filtro che è la lettura critica dell'autore stesso, dove, come osserva Enrico Iachello, «il successo dell'immagine, il consolidarsi delle descrizioni in *topoi*, nasce anche qui dall'incontro della visione "locale" con le "curiosità/aspettative" esterne»[39].

[39] E. Iachello, *La città del vulcano: Immagini di Catania* in *Catania. La Città, la sua Storia*, cit.

Lontana dall'intento di *laudatio* cartografica[40], la chiave di lettura è il confronto delle aspettative con la realtà del territorio da parte delle *élites* europee, guidate quasi sempre dai nobili autoctoni (nel caso di Houel dal principe di Biscari) che garantivano sicurezza durante il viaggio, ospitalità e mezzi di trasporto.

Durante il tour, il viaggiatore osservava con il proprio sguardo, con la sua prospettiva filtrata dall'interesse, dalle aspettative . Spesso, al ritorno nel proprio paese, «prima di scrivere le sue memorie, subiva spesso l'influenza di ulteriori letture e dell'eventuale rafforzamento di alcuni topoi, secondo la tradizione retorica, accentuando così le differenze fra il proprio paese e quello straniero, con l'attribuzione dei vizi e delle virtù ritenute caratteristiche della nazione visitata».

Nella disponibilità delle tavole digitalizzate, si è proceduto decidendo di effettuare una raccolta autonoma di tutte le tavole dei volumi, per creare un unico file PDF di illustrazioni. Questa copia garantisce un percorso di fruizione diverso e slegato dal racconto, dalla parola scritta. Le 264 tavole, ordinate progressivamente ma fisicamente divise nei 4 tomi, sono state riunite in un'unica collezione digitale.

---

[40] Contrapponendosi all'intento celebrativo dei "ritratti" di città affidati in genere da una committenza per costruirne l'identità: «Nell'ideologia e nella coscienza storica dei contemporanei, così come nelle intenzioni della committenza, questi testi figurati assumono valore analogo a quello della *Laudatio* rinascimentale: testo volto ad esaltare la storia, i commerci, la ricchezza della città figurata». C. De Seta (a cura di), *L'immagine delle città italiane dall'XI al XIX secolo*, Edizioni De Luca, Milano, 1998, pp.7-8.

14.9 Metadati

A completamento di ogni singolo tomo, si è proceduto ad associare i metadati fondamentali. Oltre al set di metadati EXIF acquisito in forma automatica dal dispositivo di scansione (*ImageWidth*, *ImageHeight*, *XResolution*, *FileSize*, *CreationDate* e *ModifyDate*), tutti i metadati sono stati inseriti attraverso la specifica XMP, *eXtensible Metadata Platform*, un formato di gestione di XML nei software Adobe che rispetta lo standard *Dublin Core*. L'inserimento è avvenuto in *batch* con l'ausilio del software Adobe Photoshop CS5.

Il set di metadati comprende *Document*, *Title*, *Author*, *Description*, *Description Writer*, *Keywords*, *Copyright Status* e *Copyright Notice*.

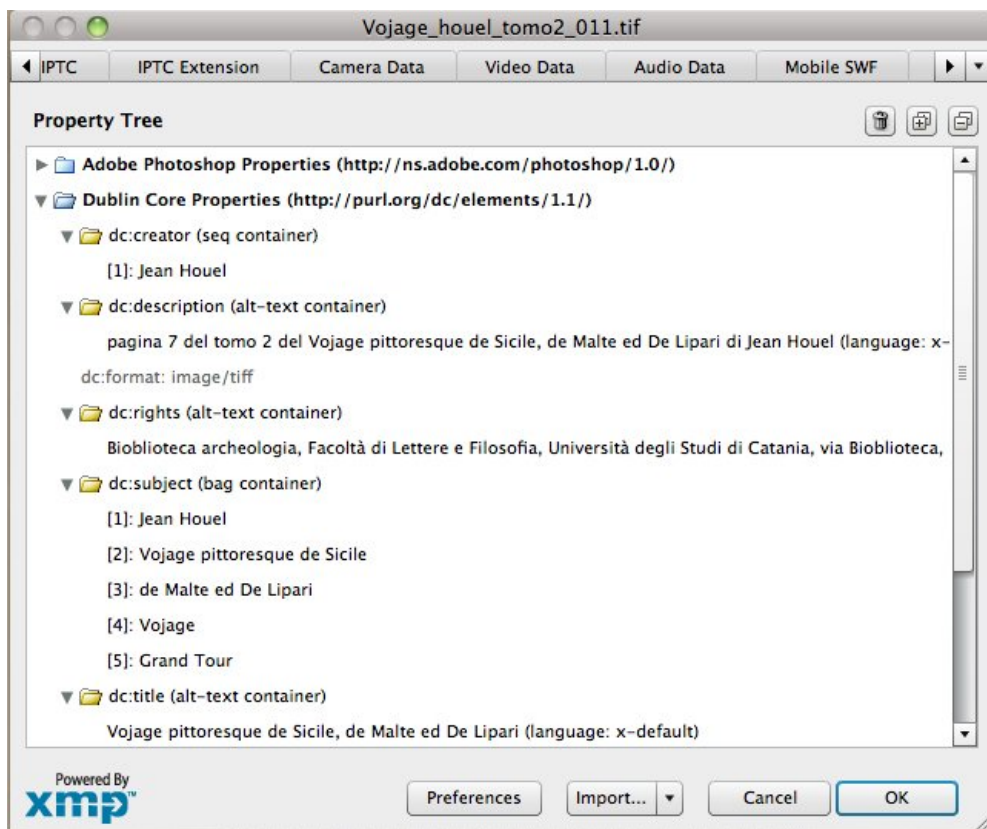



Figura 14.14.45: Una schermata di Adobe Photoshop che mostra il riepilogo dei metadati Dublin Core inseriti

14.10 Archiviazione

La prima archiviazione è stata effettuata su un Hard Disk di rete della capienza di 4 terabyte. Successivamente, l'archiviazione dei file master ad altissima risoluzione, insieme alle copie per la consultazione, in triplice copia, è stata realizzata su hard disk *FireWire*, ubicati in differenti locazioni per garantire una conservazione sicura.

14.11 Possibili sviluppi

Tra gli scenari possibili si profilano sia la possibilità di un intero restauro digitale dei volumi, che permetterebbe di ottenere il materiale già pronto per una ristampa anastatica, sia la possibilità di mettere "in rete" i files digitali, in modo tale da renderli disponibili alla consultazione da parte della comunità scientifica anche a grandi distanze fisiche grazie a internet. Sarà certamente curata la scelta di uniformarsi a standard internazionali biblioteconomici per i volumi e le biblioteche digitali, affidandosi alle struttura di metadati Dublin Core, assicurando l'interoperabilità con altre banche dati digitali a livello nazionale e internazionale.

15. La digitalizzazione del fondo De Roberto

15.1 Digitalizzare un manoscritto

Definire la peculiarità di un manoscritto in quanto oggetto digitalizzabile è un'analisi che solo a livello superficiale può apparire semplice o intuitiva. Il (*liber*) *manu scriptus*, ovvero il documento scritto a mano, in realtà è un concetto che solo a partire dall'invenzione della stampa può opporsi a quello di libro scritto o stampato in altro modo[41]. In questo senso, definire il manoscritto come un libro scritto interamente a mano, nell'accezione che comunemente oggi è condivisa, è molto riduttivo.

Scomponendolo idealmente, il manoscritto risulta essere costituito da due parti:

1. Supporto
2. Testo

[41] Come osserva Alfonso D'agostino alla voce *Manoscritto* del Manuale enciclopedico della Bibliofilia, Edizioni Sylvestre Bonnard, 1997, pp. 424-429: «A riprova, le parole *manoscritto* e, ad es., il fr. *manuscript* o lo sp. *manuscrito*, non sono attestate prima del XVI-XVII secolo. E tuttavia una certa ambiguità permane ancor oggi, se diciamo *dattilografare* o *dactylographier* (alla lettera 'tracciar segni con le dita') per quell'operazione che, facendo ricorso a un mezzo meccanico, altrove si chiama *mecanografiar*, *typewrite* etc. È che il circuito della scrittura richiede, dopo gli impulsi cerebrali, un agente [...], un oggetto su cui scrivere e un oggetto con cui scrivere o tracciare segni (ché il greco *grápho* vale appunto per entrambi i concetti). Così si produce il ms., che è la più semplice macchina di memoria; e prima della stampa è anche l'unico mezzo di riproduzione e di conservazione della cultura scritta».

Il supporto scrittorio a sua volta si distingue in supporto fisico (a livello materiale e morfologico) e strumenti grafici.

Il testo comprende invece il testo discorsivo (il contenuto trascrivibile su qualunque altro supporto), il testo decorativo (che comprende tutti gli elementi di *editing* utilizzati[42]) e, ove presente, il testo figurativo[43]. Ognuno di questi elementi è un elemento di interesse che può essere valorizzato dalla pratica della digitalizzazione, studiato e analizzato separatamente.

Altre due caratteristiche fanno di un manoscritto un oggetto fondamentale da conservare e digitalizzare: il manoscritto è *autografo* se l'esemplare è redatto di proprio pugno dall'autore e *olografo* se il documento è datato e firmato.

In quanto *autografe*, le carte antiche e scritte di pugno dagli autori costituiscono una preziosissima documentazione. Ma in quanto *olografo,* ogni manoscritto è un esemplare unico e raro, datato e fissato in determinato momento temporale, e rappresenta quindi un patrimonio fondamentale, non solo per gli studi e le analisi critiche letterarie in senso stretto, ma anche per una più ampia cornice storica che permette di inquadrare le tematiche più significative di un periodo preciso.

L'accesso ai contenuti storici e filologici del libro antico per mezzo della rete e la conservazione della documentazione mediante

[42] Si fa riferimento in questo caso a tutti gli elementi di editing, fatta esclusione delle illustrazioni, quindi: disposizione del testo, colophon, tipo di grafia ecc.

[43] Il testo figurativo è in realtà un elemento presente esclusivamente nei manoscritti che includano illustrazioni. In questo caso, rientrano in questo tipo di testo elementi disegni e miniature.

supporti tecnologicamente avanzati garantiscono la possibilità di studio e ne ampliano i campi d'indagine. Il risultato della digitalizzazione e della diffusione di un fondo manoscritto è un "ambiente" nel quale convivono differenti media e supporti (parole, immagini, informazioni, notizie bibliografiche e oggetti). Questa forma di "unità documentaria" crea un enorme ipertesto multimediale, fruibile dai luoghi fisici più diversi, in grado di fornire informazioni più disparate e, mediante ulteriori collegamenti ad altri materiali esistenti nella rete, di andare a costituire una varietà infinita di percorsi personalizzabili.

La possibilità di costruire un "catalogo visivo", di scorrere con i propri occhi le righe e gli inchiostri autografi dell'autore, consente agli studiosi una fruizione dei documenti nuova, legata a un'esperienza diretta e scrupolosa con le carte, sebbene il mezzo digitale ne costituisca un'interfaccia di mediazione.

15.2 Il progetto di digitalizzazione del fondo manoscritto De Roberto della Società di Storia Patria per la Sicilia Orientale

La digitalizzazione del fondo De Roberto fa parte del progetto *L'officina verista*, che si propone di servirsi delle più attuali tecniche di digitalizzazione per i manoscritti per avviare e

potenziare studi analitici pluridisciplinari su un autore, la cui importanza, al fine della comprensione dei processi di rinnovamento della narrativa italiana moderna in chiave storica, letteraria-stilistica e linguistico-testuale, è a tutti nota.

Nel contesto più ampio dei progetti di digitalizzazione di fondi manoscritti degli scrittori veristi siciliani, il fondo De Roberto si affianca agli attuali il processi di inventariazione e di studio dei contenuti del fondo verghiano, in buono stadio di avanzamento, come di quello capuaniano. Il fondo De Roberto è invece largamente inesplorato e, considerata l'importanza dell'autore nel quadro della cosiddetta *officina verista*, sia da un punto di vista generazionale e storico, che stilistico-linguistico, si è pensato di focalizzare l'attenzione sui manoscritti derobertiani. La digitalizzazione è stata curata dal laboratorio multimediale di sperimentazione audiovisiva la.mu.s.a. della Facoltà di Lettere e Filosofia dell'Università degli Studi di Catania, che ha gestito le fasi del progetto sia in termini ricerca e produzione, grazie alle diverse competenze dei responsabili dei settori di grafica e editoria multimediale, ma anche in termini di formazione, organizzando laboratori didattici indirizzati agli studenti della Facoltà.

15.3 Lo stato del fondo e la catalogazione

Il fondo manoscritto De Roberto è conservato presso la Biblioteca della Società di Storia Patria per la Sicilia Orientale, costituita nel marzo del 1903 dal prof. Vincenzo Casagrandi, ordinario di Storia Antica dell'Ateneo Catanese, con il fine di promuovere «lo studio della storia della Sicilia sotto tutti gli aspetti e rapporti, con speciale riferimento alle province orientali»[44]. L'istituto di Storia Patria è oggi ubicato nel centro storico di Catania, in piazza Stesicoro 29.

Il fondo De Roberto è stato costituito a partire dagli anni Ottanta, per opera del prof. Francesco Branciforti, che ha rilevato il materiale dagli eredi in vita dell'autore. I documenti non sono ancora stati inventariati, né catalogati, ma soltanto riordinati dallo stesso Branciforti in cinquantadue buste numerate, che riportano l'indicazione del contenuto. Il materiale cartaceo è composto da manoscritti, dattiloscritti, bozze di stampa, articoli di giornale, carteggi, carte geografiche, ritagli, appunti e fotografie, per un totale di circa 10.000 documenti (in appendice è riportata la lista di consistenza del fondo).

[44] C. Naselli, *Cenno storico sulle sezioni di Catania e di Messina*, in F. Brancato e R. Scaglione Guccione, *La Società siciliana per la storia patria. Storia e cultura. 1923-1993*, Palermo, 1994, pagg. 145-151.

15.3.1 Caratteristiche fisiche dei documenti

La maggior parte dei documenti sono a inchiostro bruno su carta, in dimensioni variabili dall'A5 all'A3, insieme a documenti di formati particolari (7x7 cm, 3x15cm ecc). Molti fogli hanno aggiunte e note laterali incollate a margine e ripiegate, che determinano una dimensione totale di superficie superiore all'A3. I documenti riportano numerosi interventi di correzione, tagli al testo delle opere e segni di rimando.

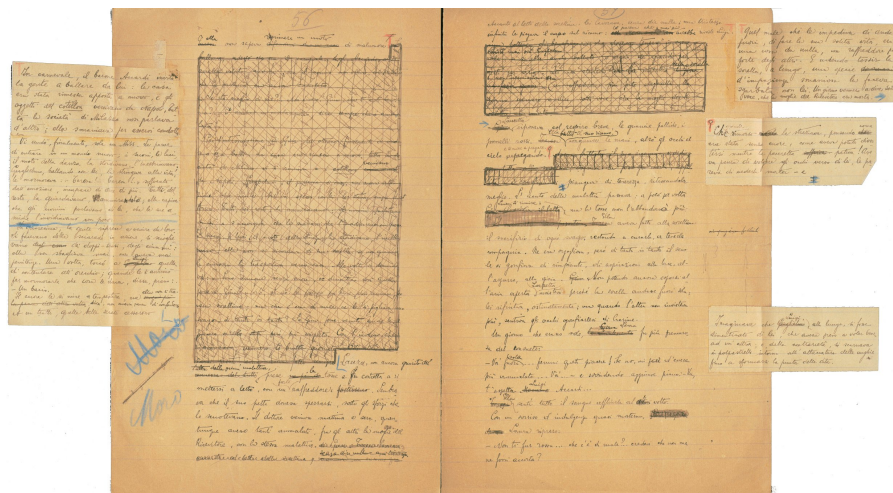



Figura 17.15.15.46: pagine 56 e 57 della prima parte del manoscritto de *L'Illusione*.

Il materiale, custodito in buste di carta, contenute a loro volta in buste portaprogetto in materiale plastico, è collocato in schedari metallici, in ambienti poco idonei alla conservazione, a causa dell'escursione termica e della forte umidità. Nonostante ciò, la maggior parte dei documenti si presenta in buono stato di

conservazione, con poche eccezioni, tra cui i preziosi appunti manoscritti per il romanzo *I Viceré*.

Figura 17.47: pagine degli appunti manoscritti per la stesura de *I Viceré*

Come è possibile vedere nella figura, il manoscritto risulta notevolmente danneggiato alle estremità, che tendono a sfaldarsi a causa della carta danneggiata da processi chimici ormai irreversibili. Per tali ragioni non è stato ancora digitalizzato, in attesa di un intervento mirato di restauro.

15.4 L'acquisizione delle immagini e il laboratorio *Manoscritti digitali*

L'acquisizione delle immagini digitali dai manoscritti del fondo è stata realizzata attraverso un progetto di formazione rivolto agli studenti della Facoltà di Lettere e Filosofia, messo in pratica attraverso il laboratorio didattico *Manoscritti Digitali*. Il laboratorio è stato curato e tenuto dalla dott.ssa Stefania Iannizzotto, dottore di ricerca e assegnista della Facoltà di Lettere e Filosofia, responsabile del settore di editoria multimediale del laboratorio la.mu.s.a.

Il laboratorio didattico ha avuto come obiettivo quello di introdurre gli studenti alle problematiche teoriche e tecniche della digitalizzazione del patrimonio manoscritto, mettendo in campo le competenze acquisite nella concreta digitalizzazione dei documenti del fondo.

Il progetto, e quindi l'attività di digitalizzazione, è stato condotto nei locali del laboratorio multimediale con le attrezzature in dotazione. Sono stati impiegati tre computer portatili e due desktop, due scanner a ripresa in piano in formato A3 e tre in formato A4, due hard disk esterni dalla capacità di un *terabyte* ciascuno.

Sono stati scrupolosamente rispettati standard e criteri di digitalizzazione conformi alle *Linee guida* e alle normative fissate dal Ministero. La modalità organizzativa e tecnica di acquisizione delle immagini è stata condotta con particolare riguardo alla qualità della

scansione, ai formati per l'archiviazione, alle convenzioni per i nomi dei file e in generale all'alta qualità dei file digitali.

I file prodotti sono stati memorizzati seguendo una nomenclatura stabilita, che prevede un codice univoco, composto da caratteri numerici e alfanumerici, che permette di riferirsi alla pagina o alla parte originale del manoscritto. Il primo numero (nell'esempio a fianco *009*) serve a tenere in ordine progressivo le immagini digitalizzate, consentendo di risalire anche alla pagina del documento fisico. Segue il titolo completo o, se troppo lungo, abbreviato dell'opera (nell'esempio *La bella morte*). Per i documenti privi di titolo, è stato brevemente riportato l'incipit del testo, in massimo quattro parole (per esempio, nella busta 48 è contenuto uno scritto non identificato, denominato, a partire dall'incipit, *Averla data a te*). Alla denominazione è stata aggiunta l'indicazione *recto* o *verso* qualora il testo fosse distribuito su entrambe le facciate.

009 la bella morte recto.tif

Figura 17.2: esempio di nomenclatura di un file

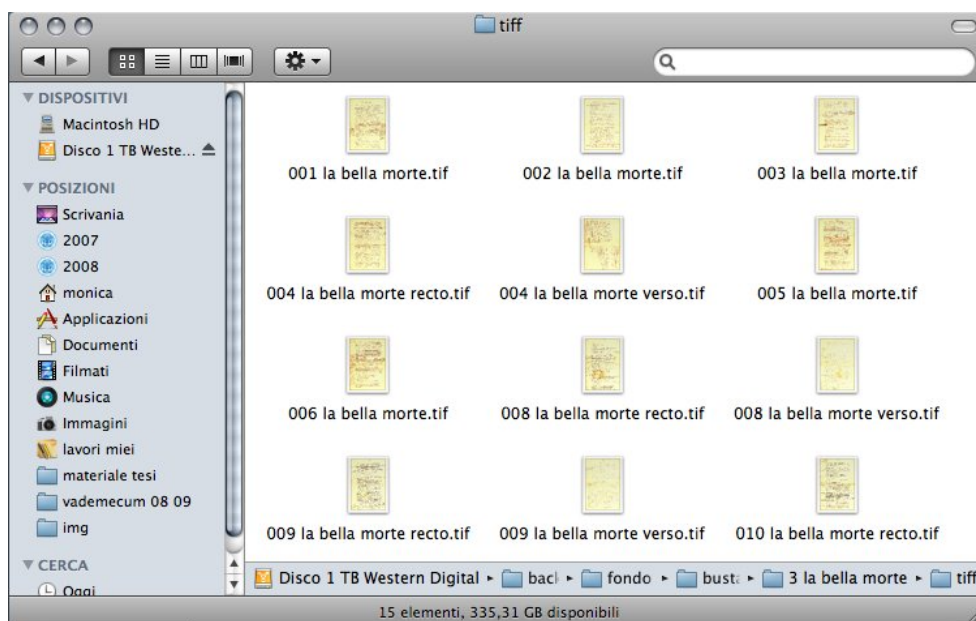



Figura 17.48: schermata che riproduce i files del manoscritto *La bella morte*, nominati secondo i criteri di nomenclatura adottati.

I requisiti di scansione adottati, quali la risoluzione, le dimensioni, la profondità di colore e i formati di output, sono quelli individuati tramite l'adattamento della *Normativa sull'acquisizione digitale*, redatta dall'ICCD nel '98, agli attuali standard *de facto*. I files master, ad altissima risoluzione (600 dpi) sono stati esportati in formato TIFF.

A completamento di ogni signola unità cartacea del fondo, rappresentata da un'opera completa o un gruppo di scritti della stessa tipologia testuale (ad esempio, *articoli politici*), si è proceduto ad associare i metadati fondamentali. Oltre al set acquisito in forma automatica dal dispositivo di scansione (*ImageWidth*, *ImageHeight*, *XResolution*, *FileSize*, *CreationDate* e *ModifyDate*), tutti i metadati sono stati inseriti attraverso la specifica XMP, *eXtensible Metadata Platform*,

un formato di gestione di XML nei software Adobe che rispetta lo standard *Dublin Core*. L'inserimento è avvenuto in *batch* con l'ausilio del software Adobe Photoshop CS.

Il set di metadati comprende *Document*, *Title*, *Author*, *Description*, *Description Writer*, *Keywords*, *Copyright Status* e *Copyright Notice*.

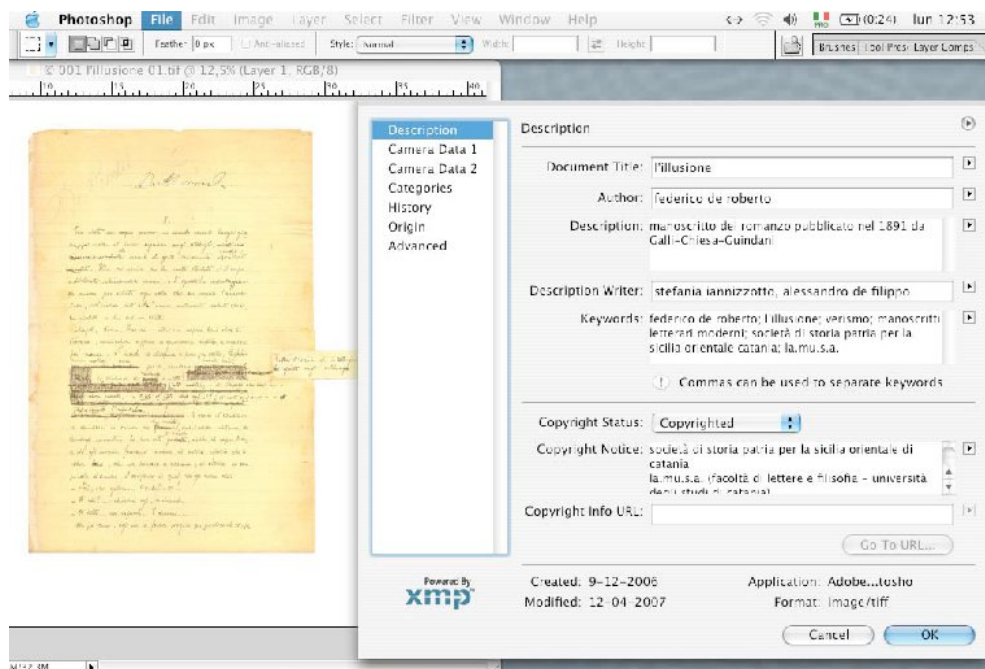



Figura 17.49: interfaccia di inserimento metadati XMP di Adobe Photoshop CS per una pagina del manoscritto de *L'Illusione*.

L'archiviazione dei files ad altissima risoluzione, in triplice copia, è stata realizzata su hard disk *FireWire*, ubicati in differenti locazioni per garantire una conservazione sicura.

15.5 Il controllo qualità sulle immagini e l'editing digitale. Il laboratorio *Elementi di fotoritocco*

A conclusione del processo di acquisizione e memorizzazione del files master, si è proceduto con un primo controllo qualitativo sulle immagini e l'*editing* e il ritocco sulle stesse. Queste attività sono state realizzate attraverso un progetto di formazione rivolto agli studenti della Facoltà di Lettere e Filosofia, che si è realizzato attraverso il laboratorio didattico *Elementi di fotoritocco*, tenuto da chi scrive in qualità di responsabile del settore di grafica del laboratorio multimediale la.mu.s.a. della Facoltà di Lettere e Filosofia.

Il laboratorio, indirizzato a un numero ristretto di studenti con competenze informatiche di base, ha avuto come obiettivo quello di fornire elementi fondamentali di *computergrafica* e ritocco fotografico, mirati a compiere operazioni di *editing* sulle riproduzioni digitali di materiale cartaceo manoscritto e dattiloscritto di fine Ottocento e inizi Novecento. Le competenza acquisite dai partecipanti hanno reso possibile il trattamento delle immagini al fine di correggere errori di acquisizione e migliorare la resa della copia digitale, conservando la massima fedeltà agli originali.

I files master del fondo manoscritto De Roberto, acquisiti duramente il laboratorio *Manoscritti Digitali*, sono stati suddivisi e assegnati a gruppi composti da due o tre studenti ciascuno, che si sono impegnati inizialmente nel controllo qualitativo delle immagini

digitali. L'analisi ha permesso di escludere l'eventualità di dover ridigitalizzare i documenti del fondo, ma ha rilevato l'esistenza di diverse unità cartacee alle quali è stato necessario applicare interventi di *editing* per migliorarne la resa qualitativa.

Una volta individuati i documenti da lavorare, sono state predisposte postazioni con dispositivi opportunamente calibrati e adatti a compiere operazioni di fotoritocco. Le immagini acquisite con bordi e margini in eccesso sono state opportunamente ritagliate e ruotate in modo tale da raddrizzarle in maniera precisa. Si è provveduto a regolare i valori di luminosità dei documenti in modo tale da migliorarne il contrasto e la leggibilità, migliorare la nitidezza dei tratti d'inchiostro e bilanciare le tonalità dei colori dei files, cercando avvicinare quanto più possibile la resa cromatica digitale ai colori originali della carta, dell'inchiostro e di eventuali correzioni e segnature.

Una volta eseguito il controllo di qualità e completate le operazioni di *editing* sui files master, sono state avviate le operazioni di esportazione per i livelli qualitativi adatti alla consultazione in sede e online (rispettivamente, alta e bassa risoluzione) e per l'associazione ai record catalografici (miniature), secondo gli standard di cui si riporta a seguire una tabella riassuntiva.

Tipologia copia	impiego	formato	risoluzione	dimensioni	prof. colore
alta risoluzione	consultazione in sede	JPEG 2000	300 dpi	originali	24 bbp
bassa risoluzione	consultazione online	JPEG	72 dpi	originali	24 bbp
miniatura	record catalografici	JPEG	72 dpi	150x150 px	16 bit

Tabella 17.21: tabella riassuntiva dei livelli qualitativi

Le immagini esportate sono state lavorate fino alla creazione di files PDF in grado di ordinarle e organizzarle in pagine, comprimendole e rendendole pronte per l'inserimento in archivi digitali.

15.6 La condivisione e la pubblicazione

Una volta portate a compimento le fasi vere e proprie di acquisizione, si è posto il problema di trovare un metodo per rendere disponibili online le immagini digitalizzate, in modo da garantirne l'accesso secondo una predefinita strategia di fruizione.

Si è ritenuto opportuno trovare un sistema in grado di garantire la disponibilità continua in rete dei documenti e la consultazione da ogni parte del mondo attraverso un'interfaccia semplice e intuitiva.

Per queste e altre ragioni, l'archiviazione online è stata eseguita sulla moderna infrastruttura grid e il supporto

dell'applicazione gLibrary. La flessibilità e l'estensibilità offerti da questo tipo di sistema hanno consentito di realizzare una perfetta integrazione nell'archiviazione e nella catalogazione in rete per i manoscritti digitalizzati del fondo. Attraverso la collaborazione con l'INFN e il Consorzio Cometa, è stato così realizzato il *De Roberto digital repository* (DRdr), di cui è disponibile una versione beta all'indirizzo <https://glibrary.ct.infn.it/deroberto>.

Il progetto è stato presentato per la prima volta il 16 giugno 2008 a Catania, durante il *Grid Open Day per l'arte, la musica e le Scienze umanistiche*[45] organizzato dal Consorzio Cometa. È stato in seguito esposto alla conferenza *eResearch Australasia 2008*[46], tenuta a Melbourne tra il 29 settembre e il 3 ottobre, nell'ambito del workshop dedicato all'*e-Research in the Arts, Humanities and Cultural Heritage*[47] e a San Miniato al *First International Workshop Cyberarch*[48] con il tema *Cyberinfrastructures and Archaeology*.

15.6.1 L'upload dei documenti e l'estrazione dei metadati

L'upload dei documenti digitalizzati è stato effettuato da utenti con permessi particolari, accordati dal complesso e sicuro sistema di autorizzazione. Attraverso l'interfaccia upload di gLibrary, che consente di selezionare files multipli, le immagini

[45] <http://indico.ct.infn.it/conferenceOtherViews.py?view=standard&confId=56>.
[46] <http://www.eresearch.edu.au>.
[47] <http://www.eresearch.edu.au/ahch-workshop>.
[48] <http://www.archaeogrid.it/cyberarch/cyberarch.html>.

digitali ad altissima risoluzione dei manoscritti e i pdf a bassa risoluzione, sono stati caricati in automatico sugli storage elements dell'infrastruttura grid Cometa.

Il procedimento di upload attraverso il *front-end* di gLibrary ha inoltre consentito l'estrazione automatica dei metadati già implementati nelle immagini, salvandoli sul server per i metadati AMGA, anche in questo caso, della grid di Cometa.

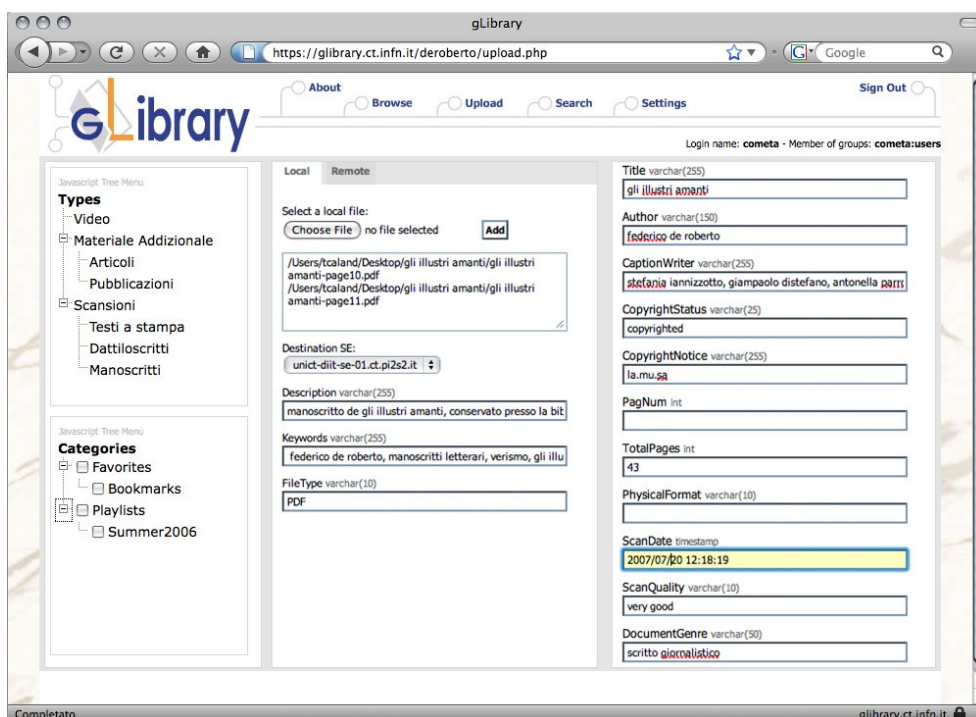



Figura 18.50: il front end di gLibrary durante l'upload dei files del manoscritto *Gli illustri amanti* di Federico De Roberto.

15.6.2 La navigazione e la ricerca

L'utente, una volta ottenuto il certificato rilasciato dalla INFN *certification authority* ed essersi registrato alla *virtual organization*

di appartenenza, effettua il *login* e accede all'interfaccia di navigazione.

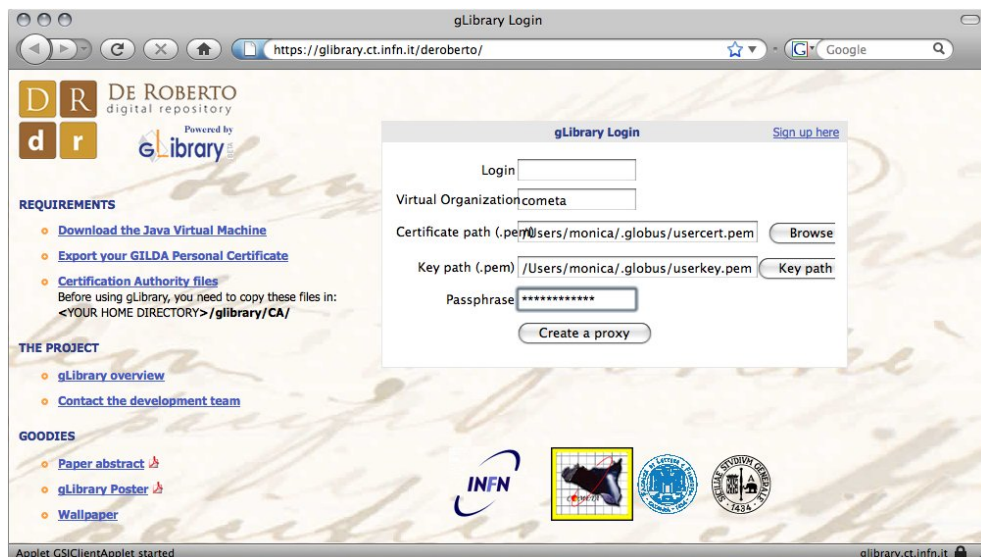



Figura 18.51: l'interfaccia di login del *De Roberto digital repository*

Il *De Roberto digital repository* consente una ricerca agevole tra gli *assets* rappresentati dai manoscritti digitalizzati: i contenuti sono organizzati per caratteristiche fisiche e semantiche attraverso la definizione di tipi e categorie e i risultati sono filtrati attraverso la selezione di uno o più attributi a scelta da liste predefinite.

Figura 18.1: l'albero dei *types* del *De Roberto digital repository*

I filtri, definiti per ogni tipo, si dividono in:

16. *DocumentGenre,* il genere a cui appartiene l'opera (novella, saggio, tragedia lirica ecc);
17. *Title,* il titolo dell'opera;
18. *FileType,* il formato del file (tiff, pdf ecc);
19. *ScanQuality,* la qualità della scansione;
20. *DocumentType,* il tipo di documento (manoscritto, dattiloscritto, bozza di stampa ecc.);
21. *PublicationStatus,* lo status di pubblicazione (pubblicato, inedito ecc.);
22. *PublicationYear,* l'anno di pubblicazione;
23. *Publisher,* il nome dell'editore;
24. *Location,* l'ubicazione fisica del documento all'interno del fondo, rappresentato dal numero della busta entro il quale è contenuto.

La selezione dei filtri, realizzata a cascata, permette all'utente di muoversi agevolmente attraverso il processo di ricerca. La selezione del valore di un filtro influenza dinamicamente la lista dei valori dei filtri successivi, consentendo di raggiungere rapidamente e in maniera razionale il risultato della ricerca.

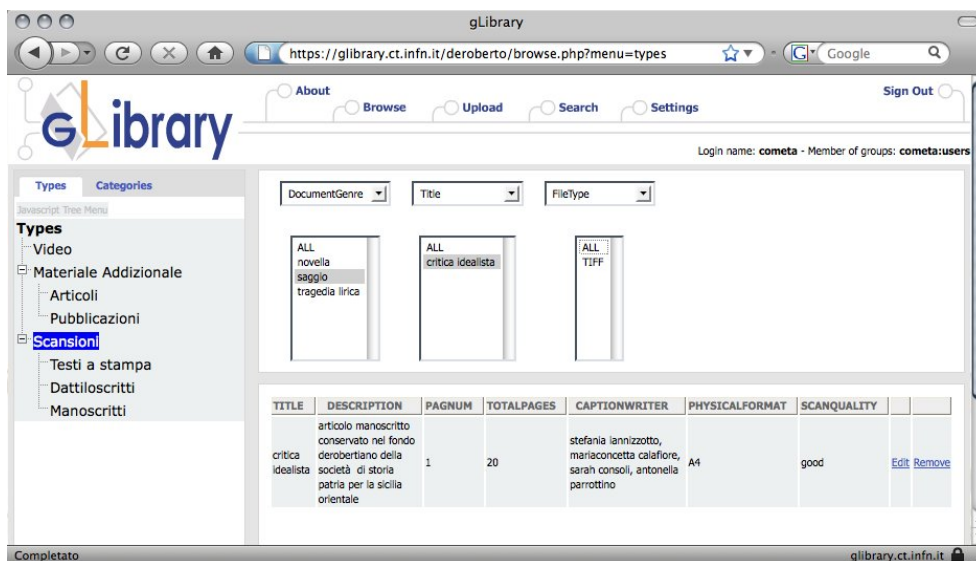



Figura 18.15.15.52: un esempio di selezione dei filtri di gLibrary sul *De Roberto digital repository* per la ricerca del manoscritto del saggio *La Critica Idealista* di Federico De Roberto

Consideriamo qualche esempio nell'utilizzo dei filtri per raggiungere un documento specifico o una serie di documenti.

Primo esempio. Uno studioso può avere bisogno di cercare, tra i documenti digitalizzati del fondo De Roberto, tutte le bozze di stampa dell'anno 1919. In questo caso, selezionerà dall'albero dei *types* disponibili *scansioni*, poi *testi a stampa*. La selezione del filtro *PublicationYear*, con il relativo valore *1919* tra gli anni disponibili, produrrà come risultato l'elenco di tutti documenti che soddisfano queste caratteristiche. La ricerca può essere ulteriormente affinata selezionando l'ulteriore filtro *Publisher*, che consentirà di visualizzare le bozze di stampa pubblicate per editore.

Secondo caso. Un linguista può aver bisogno intraprendere lo studio e l'analisi delle varianti d'autore dell'*Ermanno Raeli* di Federico De Roberto. Dovrà cercare in questo caso tutte le versioni disponibili, dal manoscritto all'opera edita. Attraverso la selezione della voce *scansioni* tra i *types* e l'impostazione del filtro *Title* con il relativo valore *Ermanno Raeli* tra i titoli delle opere disponibili, sarà in grado di visualizzare l'elenco dei manoscritti e di tutte le successive versioni delle bozze di stampa, con le relative modifiche e correzioni, dell'opera presa in considerazione, ordinabili per anno.

Il sistema, così organizzato, consente di effettuare ricerche su tutte le informazioni a livello contenutistico, ma anche dettagli sulla conservazione fisica dei documenti. È possibile ottenere risultati anche sulla qualità delle scansioni e sui formati fisici dei documenti.

È possibile accedere a una serie dettagliata di informazioni aggiuntive attraverso un semplice click sul singolo risultato, che consente di visualizzare i metadati associati ai singoli files, inseriti in fase di digitalizzazione. La gestione dei metadati è garantita dal *metadata catalogue*, il servizio che conserva e organizza i metadati dei file salvati sugli *storage elements* e registrati sul *file catalogue*.

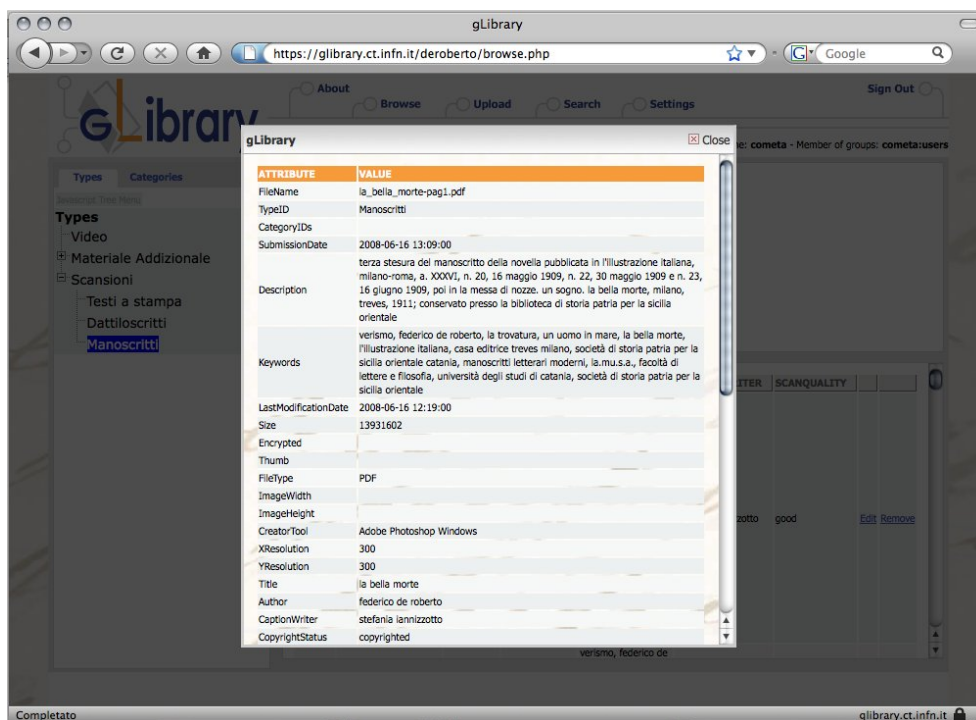



Figura 18.53: visualizzazione dei metadati associati alla prima pagina del manoscritto *La bella morte* di Federico De Roberto

Attribute	**Value**
FileName	la_bella_morte-pag1.pdf
TypeID	Manoscritti
CategoryIDs	
SubmissionDate	2008-06-16 13:09:00
Description	terza stesura del manoscritto della novella pubblicata in l'illustrazione italiana, milano-roma, a. XXXVI, n. 20, 16 maggio 1909, n. 22, 30 maggio 1909 e n. 23, 16 giugno 1909, poi in la messa di nozze. un sogno. la bella morte, milano, treves, 1911; conservato presso la biblioteca di storia patria per la sicilia orientale
Keywords	verismo, federico de roberto, la trovatura, un uomo in mare, la bella morte, l'illustrazione italiana, casa editrice treves milano, società di storia patria per la sicilia orientale catania, manoscritti letterari moderni, la.mu.s.a., facoltà di lettere e filosofia, università degli studi di catania, società di storia patria per la sicilia orientale
LastModificationDate	2008-06-16 12:19:00
Size	13931602
Encrypted	
Thumb	
FileType	PDF
ImageWidth	

ImageHeight	
CreatorTool	Adobe Photoshop Windows
XResolution	300
YResolution	300
Title	la bella morte
Author	federico de roberto
CaptionWriter	stefania iannizzotto
CopyrightStatus	copyrighted
CopyrightNotice	la.mu.s.a. - facoltà di lettere e filosofia - università degli studi di catania.società di storia patria per la sicilia orientale
PagNum	1
TotalPages	14
PhysicalFormat	
ScanDate	
ScanQuality	good
DocumentGenre	novella
DocumentType	manoscritto
PublicationYear	
PublicationStatus	published
Publisher	published
Journal	l'illustrazione italiana
JournalPublicationYear	1909
Location	busta 51
OWNER	tcaland
FILE	29
PERMISSIONS	rwx
GROUP_RIGHTS	r-x

Tabella 18.22: In dettaglio, i metadati e gli attributi contenuti nella prima pagina del manoscritto de *La bella morte*

Il download del documento desiderato può essere avviato cliccando su uno dei collegamenti disponibili rappresentati all'interno della lista delle repliche. L'archiviazione attraverso gli *storage elements* della piattaforma grid garantisce la custodia dei dati, attraverso la conservazione di copie molteplici dislocate su server in locazioni geografiche differenti, e la sicurezza nella conservazione, grazie all'affidabilità dei sistemi di storage e l'archiviazione ridondante.

I risultati delle ricerche possono essere organizzati in categorie create *ad hoc* da ogni utente, in base alle diverse esigenze, per una successiva consultazione o semplicemente per una strutturazione razionale dei documenti di interesse.

15.7 I materiali digitalizzati

- *La Lupa* – Dramma lirico in 2 atti (manoscritto)
- *Ermanno Raeli,* Milano, ed. Galli, 1889
- *I carbonari della montagna,* sceneggiatura della riduzione cinematografica (manoscritto)
- *Il convegno; Il rosario; L'onore* (manoscritto)
- *L'Illusione* (manoscritto)
- *Romanzieri italiani - G. Verga* (manoscritto)
- *Romanzieri italiani - G. Verga* (manoscritto)
- G. Verga, *Caccia al lupo* (manoscritto)
- *La tormenta* - dramma integrale corretto (manoscritto)
- *L'ebbrezza* (manoscritto)
- Verga, *La Lupa* (manoscritto)
- *Spasimo* (manoscritto)
- *La Sorte* (con indicazione "S. Giovanni La Punta - Catania 15 ottobre '82 - 12 marzo '86") (manoscritto)
- *L'Amore* (manoscritto)

- *Ermanno Raeli* (manoscritto)
- *Processi Verbali*: *Il gran rifiuto; La trovatura; Il convegno; Alla vicaria; Lupetto; Pietro Micca; L'onore; Processi verbali: Mara; Il matrimonio di Cesare; Il Rosario* (manoscritti)
- *Le confessioni*, Milano 1896
- *L'Imperio,* con correzioni autografe (dattiloscritto)
- *L'imperio*, (bozza di stampa con correzioni autografe)
- *L'Illusione* (manoscritto)
- *La prova del fuoco* (novella e abbozzo della commedia, con appunti di lavoro.
- *I Viceré (*appunti manoscritti per la stesura)
- *Il gran rifiuto* (bozza di stampa in colonna con correzioni autografe)
- *Quesiti I - II – III* (manoscritto)
- R*imorso* (manoscritto)
- *La scoperta del peccato* (bozza a stampa, strisce di carta incollata)
- *Menzogne (incompiuto) I – II* (manoscritto)
- *Il rifugio* (dattiloscritto con qualche correzione)
- *Il rosario* (dattiloscritto completo 26 pagg.)
- quaderno con sonetti a rime disegnate datato 20 novembre 1868
- traduz. datata maggio 1878 di un articolo di V. Hugo (Su Voltaire, dicembre 1823) - *Comunicazioni scientifiche e Letterarie* (articolo).

- *Voltaire* e *Comunicazioni* articoli datati maggio 1878 e novembre 1878
- Appunti sulla *Spedizione di Sicilia del 1860* (manoscritto)
- Sonetto per l'onomastico di Carolina Asmundo datato 4 ottobre 1868 firmato "l'aff.mo cognato Federico" (manoscritto)
- appunti vari per articoli di cronaca catanese (manoscritto)
- G. Verga, *La morte* Sul verso, un racconto di guerra, incompleto
- Appunti su Antonino Abate
- *L'Orgoglio e la Pietà* (2 stesure manoscritte)
- *Un proverbio italiano I-IV* (manoscritto)
- *Noi* (prima parte): 23 febbraio '91 - Michele Bardi, Corsetti, Vittorio Checchi (manoscritto)
- *Critica idealista* (articolo manoscritto)
- *Leopardi e Flaubert* (articoli manoscritti)
- *Arabeschi* (frontespizio manoscritto con indice; segue un testo su *Ricordi letterarii* di Massimo de Camp)
- *Il gran rifiuto* (manoscritto)
- *Alla Vicaria:* Nella taverna dello zio Pasquale... Bastiano [in fondo "a Federigo Casa - a F. Cameroni"]
- *I vecchi* (manoscritto)
- Bozze di lettere
- *Una voce* (manoscritto)
- *Il viaggio a San Vito* (manoscritto)

- *Il matrimonio di Cesare*: S. Giovanni la Punta 27 ottobre '89 - avv. Gilormi, cav. Placidi, cons. Collauri, Cesare Orlati (manoscritto)
- Non identificato (manoscritto pp. 10-14): Intaldi, Menfrini, Vittorio
- *Le Stagioni* I -: (bozza di stampa incollata su carta)
- Regolamenti Biblioteca
- *La retata* (manoscritto)
- *S*aggio su Leopardi (manoscritto)
- *Storia di Malta* (manoscritto)
- *Le origini di «Salammbo»* (manoscritto)
- racconto di guerra: Ciccarino di Randazzo, Gagliani
- *La delicatezza nell'arte* (3 stesure) (manoscritto)
- *L'espressione nell'arte* (2 stesure) (manoscritto)
- *Il problema del bello* (bozza di stampa) (manoscritto)
- *Analogia delle arti* (manoscritto)
- *Nel mondo ignoto* (manoscritto)
- *I destini della poesia* (manoscritto)
- minuta di una lettera a Nalli (manoscritto)
- appunti su *Vita di Verga* (manoscritto)
- appunti su *Prudhomme* (manoscritto)

- decreto di nomina a commissario della Commissione conservatrice dei monumenti e degli oggetti d'arte e di antichità per la provincia di Catania (26 febbraio 1899)
- *Il romanzo e la vita* I - II - III (su *Crimes impunis* di Marè) (manoscritto)
- *Storia dell'amore* I (manoscritto)
- *Varietà dell'amore - Il femminismo: sul libro di Bourget* (manoscritto)
- *Frammenti di novelle. L'amor supremo - L'artista - La Giustizia - Cronaca mondana - Spiegazione* (in basso a dx: Indice delle novelle che volevo scrivere) (manoscritto)
- *L'Artista* (una colonna di foglio protocollo) (manoscritto)
- *L'amore supremo* (tre quarti di colonna) (manoscritto)
- *Cronaca mondana* (tre quarti di colonna, protagonisti Guglielmo Valdara e Vico Dastri) (manoscritto)
- *Spiegazione* (manoscritto)
- *La Giustizia* (mezza colonna, prot. Landini, avv. Ruggeri, Francesco Riso) (manoscritto)
- Novelle da scrivere (una colonna): indice di quarantatre novelle (manoscritto)
- schema di personaggi per *I Viceré* (manoscritto)
- *Il memoriale del marito* (manoscritto)
- *Come siamo* (Il Male - Noi tutti, all'interno, inizi di racconti: *Noi tutti* (La Realtà); *Il Male* (manoscritto)

- *Dal profondo*: schema del racconto e inizio (manoscritto)
- Storia della proprietà letteraria del romanzo *Ermanno Raeli* con date dal 1° aprile al 18 maggio 1920 (manoscritto)
- n. 4 fotografie: Verga e i fratelli Calandra, Francofonte: municipio, già palazzo del principe di Palagonia, sede della scuola di A. Abate, folla di popolani a Francofonte.
- *Gli Amori. nuove novelle*: appunti per *Gli Amori* (manoscritto)
- Un indice per *Storia dell'amore* (manoscritto)
- *Il dramma di Venezia* IV (Teresa Lorenzo) (manoscritto)
- *Il dramma di Venezia*. Giorgio Sand. Alfredo De Musset (manoscritto)
- *Gli illustri amanti*. Giorgio Sand. Federico Chopin (manoscritto)
- *Giustizia* (manoscritto)
- Appunti su *L'Amore* (manoscritto)
- *Saggio sul libero arbitrio*: ms. non numerato in ord. perfetto: sono appunti presi da un testo, di cui si citano le pp. 282 (manoscritto)
- "Die Beit", 19 settembre 1896" - trad. tedesca di *Epilogo* di Otto Eisenschitz [prot. Carlo Landini] (manoscritto)
- *Il cane della favola* (manoscritto)
- *Nella vetrina* (manoscritto)
- appunti per articolo su *La crisi politica* (manoscritto)
- Questionario, nella colonna sx le domande, nella dx le risposte (manoscritto)

- *Una vita* (manoscritto)
- Cronaca - 1° genn. '88
- articoli vari di leggi del 1850 inerenti la giubilazione di colonnelli, luogotenenti colonnelli, maggiori
- Nuovi libri di versi
- *Romanticismo* (dattiloscritto del discorso)
- *La strada maestra,* ultimo atto[pp. 99-102] (dattiloscritto)
- Estratto da "L'Italia contemporanea", a. I, 12 ottobre 1862
- Sonetto accompagnato da una lettera a Sonzogno del 26 novembre 1868 (manoscritto)
- Spoglio del Serban (manoscritto)
- Foglietti sparsi con citazioni (manoscritto)
- Riassunto delle principali regole grammaticali francesi riunite per principianti da F.d.R. (manoscritto)
- Quaderno di citazioni in latino (manoscritto)
- Numero del Marzocco del 28 novembre 1926 (manoscritto)
- "La luce", 22 ottobre 1892 (manoscritto)
- biglietto indirizzato "All'onorevole e Distinto Cav.re Sig. Federico de Roberto Maggior comandante la Piazza in Matera" (manoscritto)
- Indice di AS: *Il serpente - Il peccato della Valcresi - Disgrazia orribile - Il marito burlone [La spada dell'angelo - Foglie al vento] - Miss -*

Convalescenza - Anna - Le Stagioni - Il rimorso - Subito - L'Enimma - Per vendetta (manoscritto)

• *La partenza dal carcere - Alla Stazione - Garibaldi - Lupetto - La Crisi - Mara - La Trovatura - Tra vecchi* (manoscritto)

• *Fatalità* (novella) pp. 1 [mancanti 2-3] – 19 (manoscritto)

• Articoli politici (manoscritto)

• Appunti per il saggio su Leopardi (manoscritto)

• Articoli politici (bozze in colonna), poi in AOL; segue l'indice del volume (manoscritto)

• Articolo su Leopardi (manoscritto)

• *Sully Prudhomme* (manoscritto)

• appunti vari (su *Prudhomme*) - fogli strappati (manoscritto)

• Articolo su *Zeusi* (Carte varie manoscritte)

• Dettagli sulla morte di Vincenzo Bellini redatto da Carlo Moncada (dattiloscritto)

• Ritaglio di giornale pubblicato sul "Giornale d'Italia" di Roma n. 179 del 29.7.1922

• Trascrizione di lettere a "Vannina"

• Fortunio (Napoli), a. III, n. 41 del 19 ott. 1890*: Le stagioni*; direttore-proprietario G.M.Scalinger

• Bozza di contratto con cui De Roberto si impegna a cedere ai f.lli Treves la proprietà letteraria di un nuovo libro intitolato *Cronache galanti* e della ristampa de *L'Illusione.*

- *Giornale di bordo* quaderno (16 pp. non numerate; solo le prime 10 pp. sono scritte) con correzioni (manoscritto)
- *Una malattia morale,* bozza di art. in colonna
- Foglietto cm.11,3 x 12,6: In questo momento penso che la saggezza consista nel rompere. 1° perché il passato non si resuscita. 2° perché sono calvo. 3° per evitare altri dolori. 4° per prendere una rivincita. 5° per essere io a rompere. [Più in basso, una "preghiera" per "Caroline", in francese con cui si invita la donna a restituire delle lettere, che accompagnano il biglietto, alla persona che le ha scritte.]
- busta (cm. 10,5 x 8,3) recante l'intestazione *Milano* e contenente un foglietto (cm. 9,7 x 15,5). Incipit: 1°. «Tu sarai l'amante della X.» Esaltazione, slancio come di gratitudine, previsione delle lacrime da spargere dinanzi a lei, e quasi pianto. - Delusione quando la vedo. Ma ripresa del lavorio della fantasia, dopo. Impazienza di conoscerla - Nuovo slancio quando l'amico dice:«Le ho promesso di presentarti»
- Foglio in colonna. Sotto l'intestazione *Milano, '94*, «Festa dell'anima, al primo incontro. Analisi: piacere della difficoltà vinta, previsione dell'esercizio dell'attività erotica, dell'ottenimento del piacere»
- Foglietto ms. cm 11 x 18, su entrambi i lati, con correzioni: «La notte dal 29 al 30 giugno '96. Ella ha detto, Signora, una parola

molto profonda: certi avvenimenti sembrano essere predisposti da un chiaroveggente destino».

• 2 bozze di lettere o di messaggi in francese con numerose correzioni.

• lettera del 31/12/97 di un cugino veneziano che si firma De Roberto;

• Gugliemino;

• Carnazza;

• Centorbi;

• Pietro di Ojieta Lanza di Trabia, lettere dell'88;

• Guido Ruberti (Giornale d'Italia); a cura di Paolo Marletta: Giselda Fojanesi; Lina Poretto De Stefano; Alessio Di Giovanni (1); Adelaide Bernardini Capuana; Sabatino Lopez

• Cesari: Piacenza, 28 ottobre '95 «Carissimo De Roberto, Il vostro libro [L'Amore] mi è giunto immensamente gradito e lo terrò come prezioso pegno della vostra amicizia.» (manoscritto)

• **L**ettera alla Società Dante alighieri (manoscritto)

• Appunti su Morny (manoscritto)

• Minute di lettere (manoscritto)

• Quadernetto di citazioni sull'amore (manoscritto)

• Partecipazione di nomina a Consigliere comunale datata 30 settembre 1910

• storia dell'elezione a consigliere comunale

- Lettere all’on. De Felice (manoscritto)
- *Il cane della favola* (per il teatro) (dattiloscritto)
- Racconto (Menfrini – Intaldi) (manoscritto)
- Voto del consiglio comunale relativo alle dimissioni (respinte) dalla carica di consigliere comunale (26 nov. 1910)
- *Una fra tante* (racconto) (manoscritto)
- *La vittima* (racconto) (manoscritto)
- Minuta di atto di compravendita terreno
- Frammento di novella. Personaggi: Rosa Maria; Evangelista Mercuri, zio Baldassarre (manoscritto)
- *Il Processo della Vita* (1-4) (frammento manoscritto di novella personaggi: Tristano, conte Alessandro; foglio con poche scritte e il titolo *La previsione dell’avvenire*; ms. *Estratto dagli «Annali di frenologia»;* ms.: *Estratto dall’Archivio di Psichiatria*, incipit: «Ero stato chiamato da poco tempo alla direzione del manicomio di *** ed avevo trovato questo stabilimento degno in tutto della reputazione che gode […]».
- *Adriana* (manoscritto pp. 35 scritte su entrambi i lati)
- Racconto inedito (52 pp.) completo datato “marzo ’84”. Incipit: «Gerolamo Lamola, a furia di eredità, aveva messo assieme una bella fortuna, da fare invidia a chiunque se non fossero state le continue liti che gli procurava» (manoscritto)
- *La bella morte*, 3 stesure, la prima intitolata *Un uomo in mare* (manoscritto)

- Il *diritto di morire - studio etico-giuridico*, con prefazione di Raffaele Garofalo, Roma, Casa Editrice Italiana, 1895 (sul suicidio), all'interno un altro libro intitolato *Del suicidio in rapporto alla morale, al diritto ed alla legislazione penale positiva*. Largo sunto delle quattro conferenze dell'avvocato L. Zuppetta professore di legislazione penale comparata nell'Università di Napoli, seconda edizione, Napoli, Ernesto Anfossi Editore, 1885
- Appunti per *La venere di Siracusa* e su città greche (manoscritto)
- *Sull'orlo* (autore ignoto) (manoscritto)
- *Taddeo e Veneranda,* novella di Edoardo Sboto (10 pp.) (dattiloscritto)
- E. Sboto, *Mors tua*...(novella) (dattiloscritto)
- "Il Giornale d'Italia", 31 dic. 1914
- *D'Urfè* (articolo Giornale d'Italia)
- *La Salvazione* (1-6) (bozza di stampa)
- *Il Paradiso perduto* (3-7) (bozza di stampa)
- Zakunine (manoscritto)
- Carta geografica militare del presidio di Catania con scritta sul retro "S. Ten. De Roberto"(di Diego, probabilmente), Istituto Geografico Militare 1892.
- *Per la musica* (manoscritto)
- *Lezioni della storia - La guerra per Malta* (manoscritto)

- *Traité de paix entre la République Francaise et le Roi des deux Siciles: signé le 28 Mars 1801* (manoscritto)

16. *I Cantieri di Giancarlo De Carlo a Catania*: un prodotto multimediale sul lavoro di recupero e restauro del Monastero dei Benedettini

16.1 Dall'editoria tradizionale a quella multimediale

In senso tradizionale, multimedialità è la possibilità di lavorare contemporaneamente con più mezzi di comunicazione. Dalla nascita delle prime tecnologie multimediali, il termine ha identificato la compresenza di almeno tre strumenti: testo, audio e immagini (sia fisse che in movimento). È significativo il fatto di poter ottenere, attraverso l'interazione di questi elementi sensoriali, un cambiamento nella comunicazione e nella interattività: poter passare, dunque, da un concetto all'altro, da un contenuto all'altro, utilizzando simultaneamente più media.

Nell'ambito dei Beni Culturali, e nello specifico nel recupero e nella valorizzazione degli stessi, un prodotto multimediale è attualmente lo strumento migliore in grado di documentare il recupero di un bene in tutte le sue fasi, avvalendosi dell'ausilio di attraverso diversi media. Si tratta di una forma nuova di legame tra

contenuti e capacità di interazione: l'utente è in grado di "esplorare" con percorso proprio, riuscendo a far dialogare e compenetrare elementi tecnici ed elementi umanistici in una forma nuova.

16.2 Le fasi di ricostruzione del Monastero dei Benedettini e il multimediale

Il vasto progetto di recupero e restauro condotto dall'architetto Giancarlo De Carlo, durato trent'anni, ha reso possibile l'adeguamento del complesso monastico di San Nicolò l'Arena come sede delle Facoltà di Lettere e Filosofia dell'Università degli Studi di Catania. Dal 1977 ad oggi, il progetto di De Carlo si è concentrato nella trasformazione di uno spazio, di un edificio già di per sé adibito a numerose funzioni, per «farlo diventare un luogo universitario, senza compiere interventi troppo forti che non avrebbe [..] tollerato[49]». Si è trattato dunque di un progetto concentrato nella rifunzionalizzazione di un edificio «che viene liberato dalle più recenti superfetazioni e riparato dalle mutilazioni subite soprattutto nei primi cinquant'anni di questo secolo».

L'opera di De Carlo e i documenti prodotti durante le fasi di realizzazione sono molto più che semplici relazioni o fotografie di cantiere: come afferma Giuseppe Giarrizzo, «De Carlo ha sempre dichiarato il problema della coerenza, cioè il problema del modo col

[49] Giancarlo De Carlo, *Un progetto per Catania. Il recupero del Monastero di San Nicolò l'Arena per l'Università*, Sagep, Genova, 1988, p.

quale una struttura fisica realizzata da un progetto si rapporta verso altre funzioni[50]». Lo stesso De Carlo ha più volte preso le distanze da un progetto di restauro di natura esclusivamente conservativa:

> […] si potrebbe pensare che il progetto sia stato prevalentemente orientato al restauro conservativo. Invece non è proprio così perché, guardando nelle articolazioni, ci si accorge di un diffuso ribaltamento dei significati ottenuto insinuandosi nelle discontinuità per ricomporle a diverse scale in modo da non alterare […] l'unità della configurazione complessiva[51].

I cantieri di Giancarlo De Carlo a Catania è un progetto che, attraverso la realizzazione di un DVD multimediale, raccoglie tutti i documenti in formato digitale del progetto di recupero del Monastero dei Benedettini in formato digitale prodotti da Giancarlo De Carlo, dai suo collaboratori e dall'ufficio tecnico.

---

[50] Giuseppe Giarrizzo, Introduzione video a *I cantieri di Giancarlo De Carlo a Catania*, in corso di stampa

[51] Giancarlo De Carlo, *Un progetto per Catania,* cit.

16.3 I contenuti

16.3.1 Una introduzione multimediale: l'intervista dei prof. G. Giarrizzo e E. Iachello

Per questo progetto multimediale è stata adottata la scelta di sostituire le introduzioni classiche dell'editoria tradizionale con introduzioni video. Sono così stati realizzati i due filmati introduttivi del prof. Giuseppe Giarrizzo, Preside per oltre trent'anni della Facoltà di Lettere e Filosofia che ha raccontato lo svolgersi dei lavori all'interno del Monastero, e dell'attuale preside Enrico Iachello.

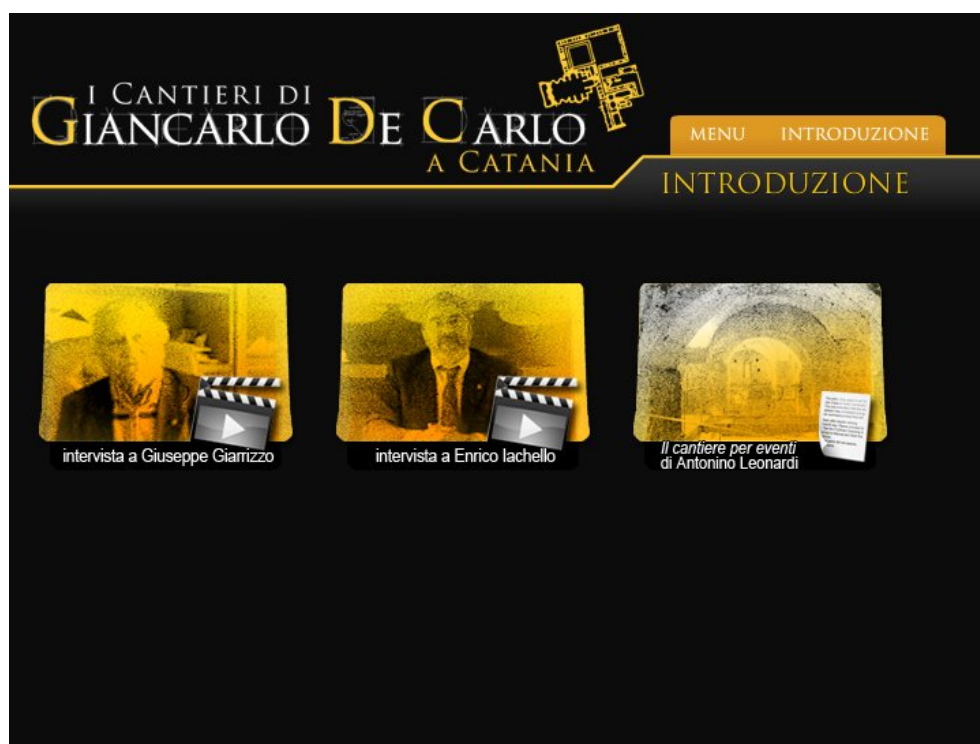



Figura 16.54: La schermata dell'introduzione de *I cantieri di Giancarlo De Carlo a Catania*

Figura 16.55: L'introduzione video del prof. Giuseppe Giarrizzo

16.4 Le relazioni di Giancarlo De Carlo

Giancarlo De Carlo ha redatto un totale di 93 relazioni di cantiere dal marzo del 1989 al gennaio del 2004. Era solito elaborare questi documenti per ogni sopralluogo che effettuava nei diversi periodi, descrivendo dettagliatamente lo stato di avanzamento dei lavori ma anche molti altri elementi, come le relazioni con i tecnici e le riflessioni personali. Il lascito testuale di De Carlo sul Monastero dei Benedettini diventa così non solo testimonianza per gli addetti ai lavori, ma anche una lettura, un racconto sui luoghi e le persone, sugli incontri e sulle attività, un

diario che prende spunto dal lavoro di cantiere per raccontare molto altro. Come ha sostenuto Giuseppe Giarrizzo,

> Chiunque si trovi con questo materiale, qualunque sia la ragione per la quale intende leggerlo, deve avere presente che questo non è un libro scritto da De Carlo, che De Carlo non scriveva queste sue relazioni avendo in prospettiva quelle di pubblicarle o di rielaborarle. Ma avrebbe potuto benissimo cosa che poiché era uno scrittore splendido[52].

Le relazioni all'interno del DVD mantengono la numerazione originale e sono state divise cronologicamente per anno per una navigazione più semplice.

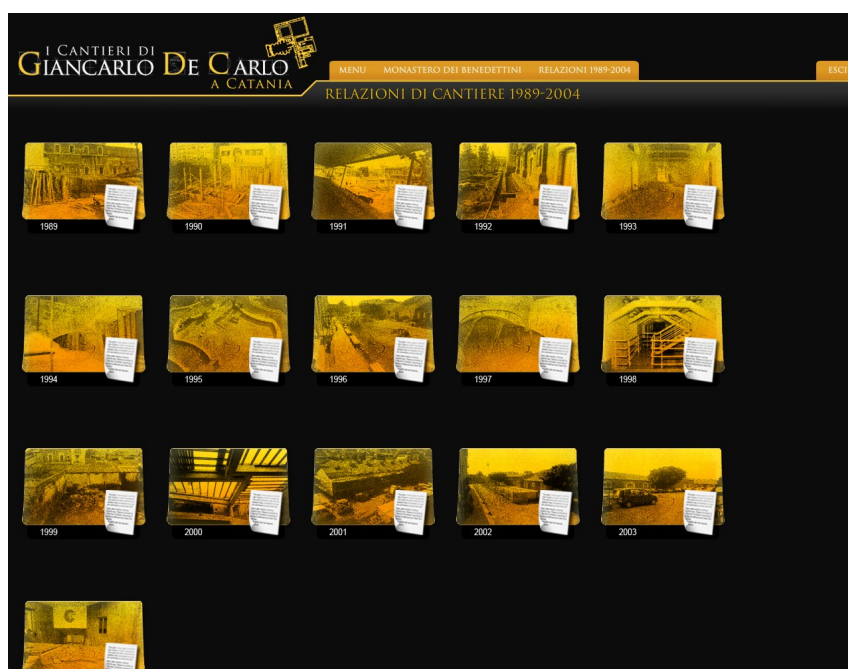



Figura 16.56: Una schermata che riproduce l'indice delle relazioni, suddivise per anno

[52] Giuseppe Giarrizzo, Introduzione a *I cantieri di Giancarlo De Carlo a Catania*, Cit

Ogni relazione contiene le immagini, sia foto che progetti, schizzi preparatori, disegni, dei luoghi trattati. I nomi degli spazi sono evidenziati per renderli immediatamente associabili alle immagini a fianco.

Figura 16.57: Una schermata che riproduce la relazione n.14 del 10 e 11 maggio 1991

16.5 Le gallerie d'immagini

Le gallerie di immagini, suddivise per anni come le relazioni, contengono le fotografie dello stato di fatto dei lavori nelle varie

fasi di avanzamento, le fasi di lavorazione e, spesso, immagini dei tecnici che hanno lavorato ai progetti.

Figura 16.58: Una schermata che riproduce l'indice delle relazioni, suddivise per anno

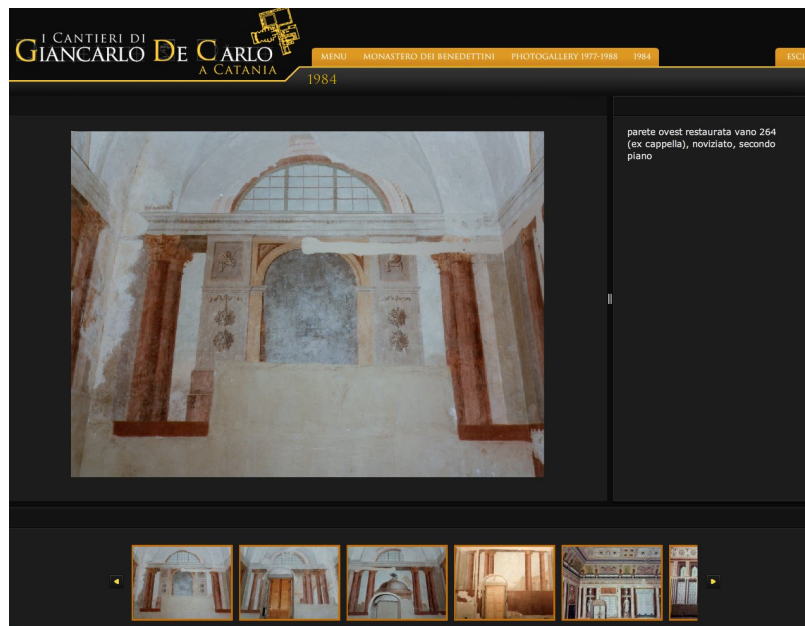



Figura 16.59: La photogallery dell'anno 1984. In basso, le miniature delle immagini selezionabili. A sinistra l'immagine selezionata, a destra la didascalia di riferimento.

Le immagini, ad alta risoluzione, sono navigabili per mezzo di una barra di zoom: l'utente può ingrandire e rimpicciolire a piacimento, osservando, se lo desidera, dettagli e particolari, diversamente dalla tradizionale stampa fotografica. Le immagini possono essere visionate nell'ordine stabilito, per mezzo dei tasti di avanzamento, o selezionate senza seguire un ordine dalla miniature in basso.

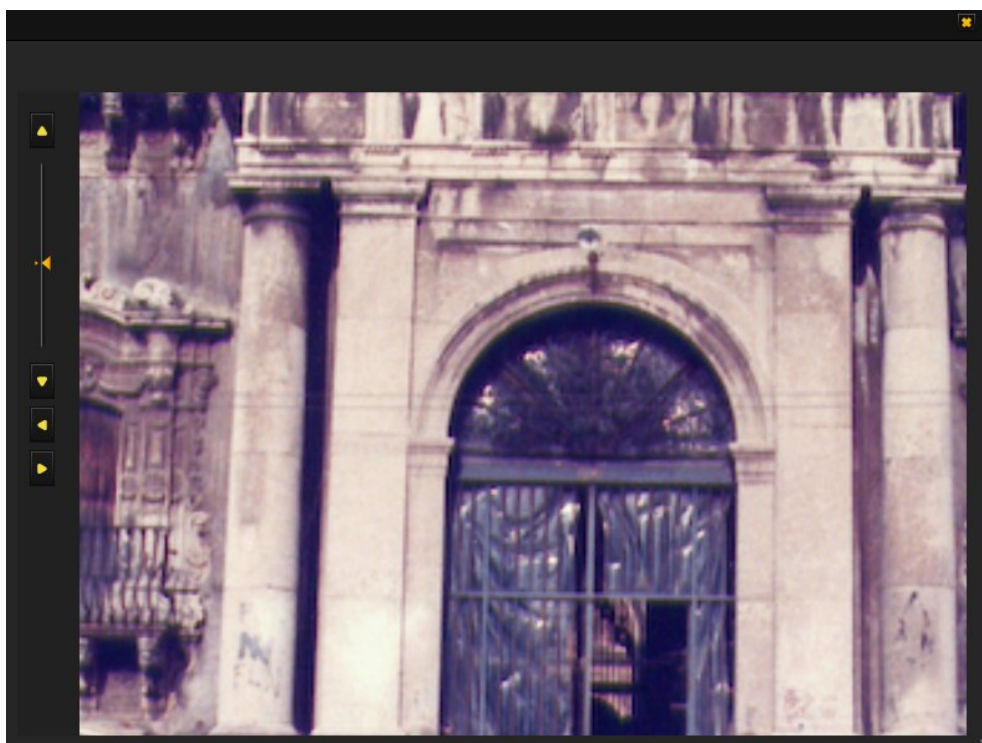



Figura 16.60: Il sistema di navigazione delle immagini: a sinistra, una barra permette di zoomare aventi o indietro, ingrandendo o rimpicciolendo. Subito sotto, i tasti per avanzare alle immagini precedenti o successive

16.6 Gli altri contenuti

Il DVD multimediale offre la possibilità di consultare anche altri tipi di contenuti, che hanno a che fare con Giancarlo De Carlo, con il Monastero dei Benedettini o con altri progetti.

Una biografia completa di Giancarlo De Carlo si affianca alle immagini dei lavori, delle opere o dei momenti dell vita dell'architetto.

Figura 16.61: La biografia di Giancarlo De Carlo

La bibliografia è suddivisa in tre parti: gli scritti di Giancarlo De Carlo, quelli su Giancarlo de Carlo e gli scritti in generale sul

Monastero dei Benedettini. I testi indicati, oltre a contenere le indicazioni bibliografiche classiche, sono accompagnati dall'immagine di copertina, visionabile anche in ingrandimento.

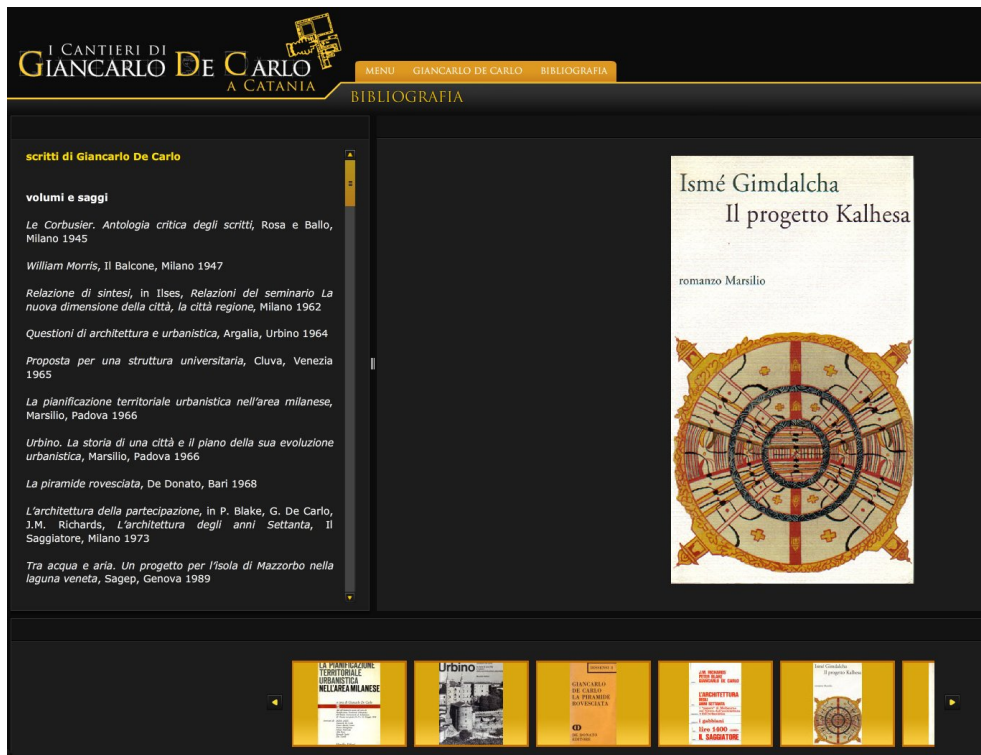



Figura 16.62: Una schermata che riproduce la bibliografia e l'ingrandimento della copertina di uno dei testi selezionati

Un'altra sezione contiene tutte le opere e i progetti di Giancarlo De Carlo realizzati in Italia e all'estero in ordine cronologico

Figura 16.16.63: La schermata delle opere e dei progetti di Giancarlo De Carlo

16.7 La struttura del dvd

1. **Introduzione**

1.1 Intervista a Giuseppe Giarrizzo

1.2 Intervista a Enrico Iachello

1.3 *Il Cantiere per eventi* di Antonino Leonardi

2. **Giancarlo De Carlo**

2.1 Biografia

2.2 Bibliografia

2.2.1 Scritti di Giancarlo De Carlo

17. *Con la sicilia negli occhi:* un documentario sul documentario siciliano

17.1 Il documentario e la sua funzione storica

Il cinema è da sempre stato un mezzo conoscere e per interpretare la realtà. Il documentario, nel suo tentativo di mostrare il reale, potrebbe essere in questi termini considerato attendibile come fonte storica, come veicolo di contenuti che testimoniano la realtà. Ma il discorso non è così semplice. La fonte, il documento, poco hanno a che fare con il documentario. Sebbene non si tratti di *fiction*, di messinscena, il documentario è sempre frutto di un filtro da parte dell'autore, del regista, delle figure che operano in qualche modo una selezione sulla realtà. Pur volendo mantenere il più severo criterio di oggettività sul reale, il cinema è un *punto di vista*, come spiega Alessandro De Filippo: «conserva il peccato originale del *punto macchina*, della posizione di ripresa dell'autore, laddove la scelta dell'osservatore prima e quella del narratore immediatamente dopo denunciano la loro inevitabile soggettività[53]».

Peppino Ortoleva riassume bene questo discorso con due realtà che definisce *sogno* e *testimonianza:*

[53]Alessandro De Filippo, *Con la sicilia negli occhi. Il cinema documentario di Ugo Saitta*, in corso di stampa, Bonanno Editore.

> "Sogno" e "testimonianza", non possono che essere due anime separate del cinema, si rinviano continuamente una all'altra. Il carattere mistificatorio del concetto di documentario come verità auto- certificantesi; la possibilità di usare il cinema, anche il più evidentemente propagandistico e falsante, come traccia preziosa di un mondo scomparso; l'ambivalenza irriducibile del cinema e della sua verità [...]. La macchina delle verità e la macchina dei sogni, anziché escludersi, si richiamano e si fondono a vicenda: il film come documento veritiero ed il film come illusione ed anche mistificazione sono due facce del fenomeno-cinema [...]. Nella moderna cultura di massa, in particolare in quella audiovisiva, reale e fantastico, documentazione e rappresentazione, si definiscono e si pongono a vicenda: con un'ambivalenza che ha le sue radici più profonde in alcune specifiche caratteristiche antropologiche dell'umanità contemporanea. È questa ambivalenza, l'intreccio di macchina della verità e macchina dei sogni, che ha permesso al cinema prima e poi alla TV di assumere il ruolo centrale che hanno avuto ed hanno nelle società contemporanee[54].

Il cinema non è «testimone infallibile». È anche «macchina dei sogni» e, in quanto tale, veicolo di contesti sociali, culturali. In questo senso, però, può avvicinarsi allo storico: permettendogli di analizzare il contesto in cui l'autore opera e divenendo, a sua volta, strumento di ricerca e di indagine.

Tutto il lavoro di Ugo Saitta, tutti i suoi documentari e tutta la sua filmografia si intrecciano con la storia della Sicilia, raccontando i cambiamenti economici, sociali, culturali dell'ultimo

[54] Peppino Ortoleva, *Testimone infallibile, macchina dei sogni: il film e il programma televisivo come fonte storica*, in «Passato e presente. Rivista di storia contemporanea», luglio-dicembre 1984.

secolo: l'invasione turistica alla fine degli anni Cinquanta, il boom economico e le industrie a metà dei Sessanta, la trasformazione delle arti e dei mestieri, le mutazioni sociali e le "rivoluzioni" del costume. Il cinema di Saitta racconta storie. Testimonianze oggettive o meno, i suoi documentari sono il suo sguardo sulla Sicilia del Novecento, testimone/macchina dei sogni che produce immagini che, oggi, possiamo utilizzare come strumento di ricerca: «partiamo dal *come* per capire il senso del *cosa* Saitta ha raccontato della Sicilia[55]».

17.2 Il cinema documentario di Saitta

Ugo Saitta (1912-1983), socialista catanese, attivo a partire già dal 1934, ha prodotto una trentina di cortometraggi documentari, un corto di animazione (*Teste di legno*, 1939) e un lungometraggio di finzione (*Lo voglio maschio*, 1971). I temi principali del suo cinema sono l'Etna, lo sviluppo economico dell'isola (anche per mezzo dell'industria, ottimisticamente desiderata e attesa negli anni del boom economico) e la ricchezza artistica della Sicilia, da Siracusa alla Valle dei Templi, da Palermo all'amatissima Catania.

Saitta realizza film molto diversi tra loro, diversi per intenzione e per ruolo comunicativo. Se *Zolfara* (1947) ci appare estremamente vicino all'esperienza neorealista e quindi anche allo

[55] Alessandro De Filippo, *Con la sicilia negli occhi,* cit.

spirito polemico e politico di quell'estetica, *L'Etna è bianco* (1947) o *Sant'Agata* (1950) sembrano avvicinarsi più alla descrizione "neutra" e puntigliosamente descrittiva della *Settimana Incom*, di cui l'autore fu anche attento e instancabile collaboratore, esteticamente più tradizionale e conservativa; i documentari *La valle dei templi. Agrigento* (1950) e *Mito e realtà di Siracusa* (1951) ci appaiono invece totalmente concentrati nello sforzo propositivo e ottimista della promozione turistica della sua amata isola. C'è un'intenzione giornalistica in *Volti di Sicilia* (1951-1969, anche nella struttura tripartita dei servizi) e c'è un grande uso delle immagini di repertorio, un archivio ricchissimo che Saitta riutilizzava in maniera sapiente. Infine l'autore catanese ricorre di frequente alla contaminazione con alcune sequenze di finzione, messe in scena e recitate da attori "caratteristi" come Tuccio Musumeci, Carmela Trovato, Riccardo Mangano, Turi Scalia e Mario Piazza. Una ricerca piena di "derive" e di commistioni linguistiche, di esperimenti espressivi e di ripensamenti mediatici, che necessita di un'indagine conoscitiva approfondita.

17.3 Le fasi del progetto: la ricerca, la catalogazione dei filmati e la stesura della filmografia completa

Il lavoro di studio e recupero sul lavoro di Ugo Saitta ha attraversato fasi diverse che hanno coinvolto numerose figure e istituzioni.

La prima parte della ricerca, di carattere esplorativo, è avvenuta grazie alla collaborazione della Filmoteca Regionale Siciliana, che ha permesso la fruizione ripetuta delle opere di Ugo Saitta, tutte conservate in formato 35mm e in digitale (su supporto DVD). La disponibilità del direttore e dei suoi collaboratori, su tutti di Marcello Alajmo, ha permesso di conoscere, quasi per intero, la produzione dell'autore siciliano: mancano infatti i lavori precedenti alla seconda guerra mondiale (a esclusione di *Teste di legno*, del 1939), che non sono nella disponibilità dell'istituzione siciliana.

La seconda fase è stata resa possibile dall'incontro con Gabriella Saitta, la figlia del regista, che con grande generosità ha messo a disposizione, per una fruizione preliminare, un album pieno di documenti e di articoli raccolti dal padre. Rispondendo alle nostre domande, ci ha dato anche delle indicazioni su possibili canali di ricerca alternativi rispetto alla famiglia, invitandoci a contattare Sebastiano Gesù, docente di Storia del cinema della Facoltà di Lingue e Letterature straniere di Catania; Roberto Lanzafame, che ha collaborato con la moglie di Saitta subito dopo la sua scomparsa, per riordinare materiali e documenti che lo riguardavano; e Sarah Zappulla Muscarà, che sarebbe in possesso di un grande numero di fotografie e documenti, donati dalla stessa Gabriella al momento della vendita della sua casa catanese (intorno al 1987, dopo la morte della madre).

La terza fase ha visto la riorganizzazione dei materiali già acquisiti in digitale (29 tra documentari e cinegiornali) e l'analisi tecnica ed estetica degli stessi; lo studio dei pochi articoli pubblicati da Sebastiano Gesù in due diversi testi editi da Maimone, *La Sicilia della memoria. Cento anni di cinema documentario nell'isola* (1999) e *L'Etna nel cinema. Un vulcano di celluloide* (2005). In questa stessa fase si è proceduto alla scelta delle *locations* per la ripresa e alla selezione dei collaboratori componenti la *troupe*. Inoltre si è proceduto alla realizzazione del piano finanziario del progetto affidando al dott. Natale Patania, assistente amministrativo della Facoltà di Lettere e Filosofia dell'Università degli Studi di Catania, la predisposizione del budget di previsione per la realizzazione del documentario.

17.4 Le immagini e i documenti e digitalizzati

17.4.1 Dall'archivio privato di Ugo Saitta

Ugo Saitta era solito raccogliere in maniera scrupolosa ogni documento che riguardasse i suoi film: le foto di scena, gli articoli di giornale sui suoi lavori, i fogli di sala delle presentazioni dei suoi lavori. Questi materiali, conservati dalla figlia del regista, Gabriella Saitta, hanno costituito un corpus importantissimo per la

ricostruzione della filmografia, della cronologia, ma anche delle informazioni biografiche sulla vita del regista.

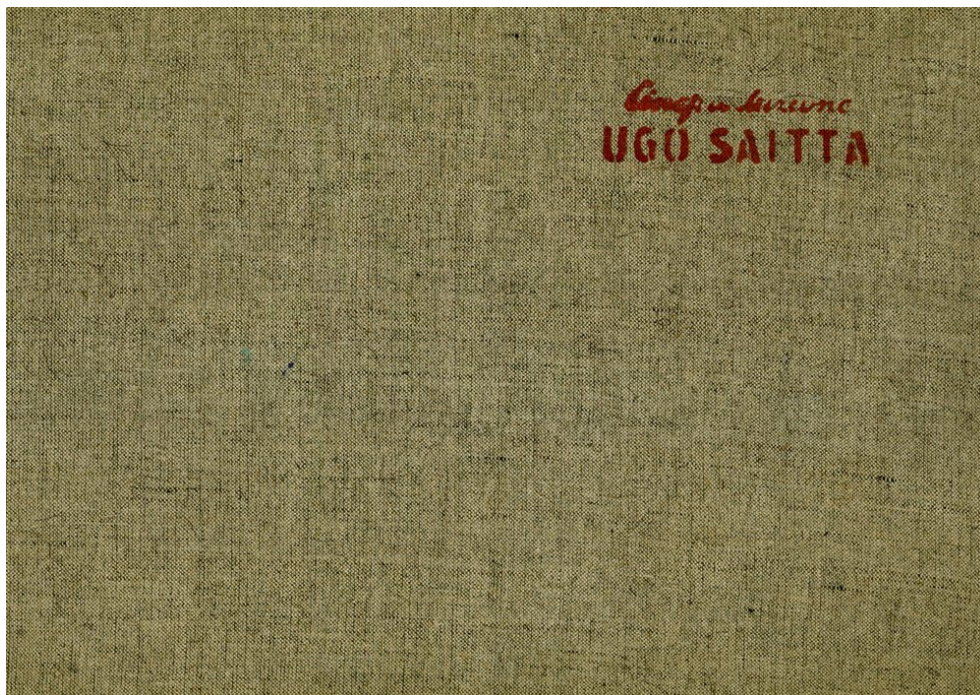



Figura 17.64: La copertina dell'album

L'album/archivio è stato interamente digitalizzato ad alta risoluzione per mezzo di scanner piani A3. Si tratta di materiali di grandi dimensioni; le fotografie e gli srticoli di giornale sono stati incollati dallo stesso Saitta con una composizione molto particolare e creativa. Sono state, pertanto, digitalizzate le pagine intere, mantenendo l’impostazione originale.

SOCIETA' CINEMATOGRAFICA ITALIANA
ADERENTE AL DOPOLAVORO CINEMATOGRAFICO ROMANO
SEDE CENTRALE

Ufficio PRESIDENZA
Prot. N. 839/V.A.
Risposta al n. = del 26/10/1934 XII

ROMA, li 5 Novembre 1934 XIII
Via Poli, 25 - Tel. 61.740

Preg.mo Sig.Ugo Saitta
Presidente della Sede Prov.
di CATANIA

Caro Saitta,

in riscontro alla sua gradita sopranotata:

Voglio sperare che dalle promesse di attività da lei fattemi e dallo zelo ed appassionato fervore che lo animano, potremo presto trarre ottimi risultati a favore della "S.C.I.".

Dato che la "S.C.I." diverrà fra breve Ente Nazionale per il Cinedilettantismo ed assorbirà tutti i cinegruppi esistenti in Italia, crede non sarebbe meglio fondere completamente la sua "Juventus Film" nella nostra Società, lasciando l'unica denominazione di "Società Cinematografica Italiana" Sede Provinciale di Catania? Sarebbe un ottimo mezzo per evitare dispersione di forze e confusioni.

Ho molto gradito l'omaggio delle fotografie del suo film "Clima puro" (ottime) ed avrò molto care quelle che mi promette d'inviare poi.

Non credo sia opportuno inviare una piccola parte del film "Clima puro"; potrebbe invece inviare "Vent'anni". Questo film desidererei anzi che lei me lo inviasse corredato, se possibile, dalle inquadrature fotografiche o plenches ad esso relitive. Le nostre manifestazioni sono state rimandate di qualche tempo dato che saranno svolte con maggior fasto in collaborazione con la Direzione Generale per il Cinematografo del Sottosegretariato per la Stampa e la Propaganda. "Clima puro" potremo presentarlo, appena pronto, in visione privata alle Gerarchie ed alla stampa e, se buono, potremmo interessarci noi stessi per piazzarlo sul mercato.

Le accludo una lettera del Sig.Giuseppe Di Mauro. La prego di esaminarla e di mandare a chiamare lo scrivente. Potrebbe divenire nostro socio. La prego provvedere per il tesseramento e relazionarmi sulla attività della Sede.

Cordiali saluti.

IL PRESIDENTE CENTRALE
(Vinicio Araldi)

SOCIETA' CINEMATOGRAFICA ITALIANA
Sede Centrale: ROMA - Via Poli, 25
Tel. 61.740

Preg.mo
Sig. UGO SAITTA
Juventus Film
Via Plebiscito,887
CATANIA

Figura 17.65: Una lettera della società cinematografica Italiana

Il Sig. Saitta
Ugo
domiciliato in Catania
fa parte della "Società Cinematografica Italiana" in qualità
di Presidente Provinciale

Firma del Titolare: Saitta Giuseppe Ugo

Il Presidente

Figura 17.66: La tessera della Società Cinematografica Italiana

Evening **SERA** Soir

1) **Les Goemons**

Documentario (*FRANCIA*)

Produzione Etienne Lallier
Musica Guy Bernard

2) **L'AIGLE A DEUX TETES**

(L'acquila a due teste) (*FRANCIA*)

Regia Jean Cocteau.
dal monologo teatrale di Jean Cocteau.
Fotografia Christian Mastras
Scenografia Georges Wakhevitch.
Musica Georges Auric.
Interpreti Edwige Feuillère, Jean Marais, Silvia Monfort, Jacques Varennes, Jean Debucourt.

3) **La Zolfara**

Documentario (*ITALIA*)

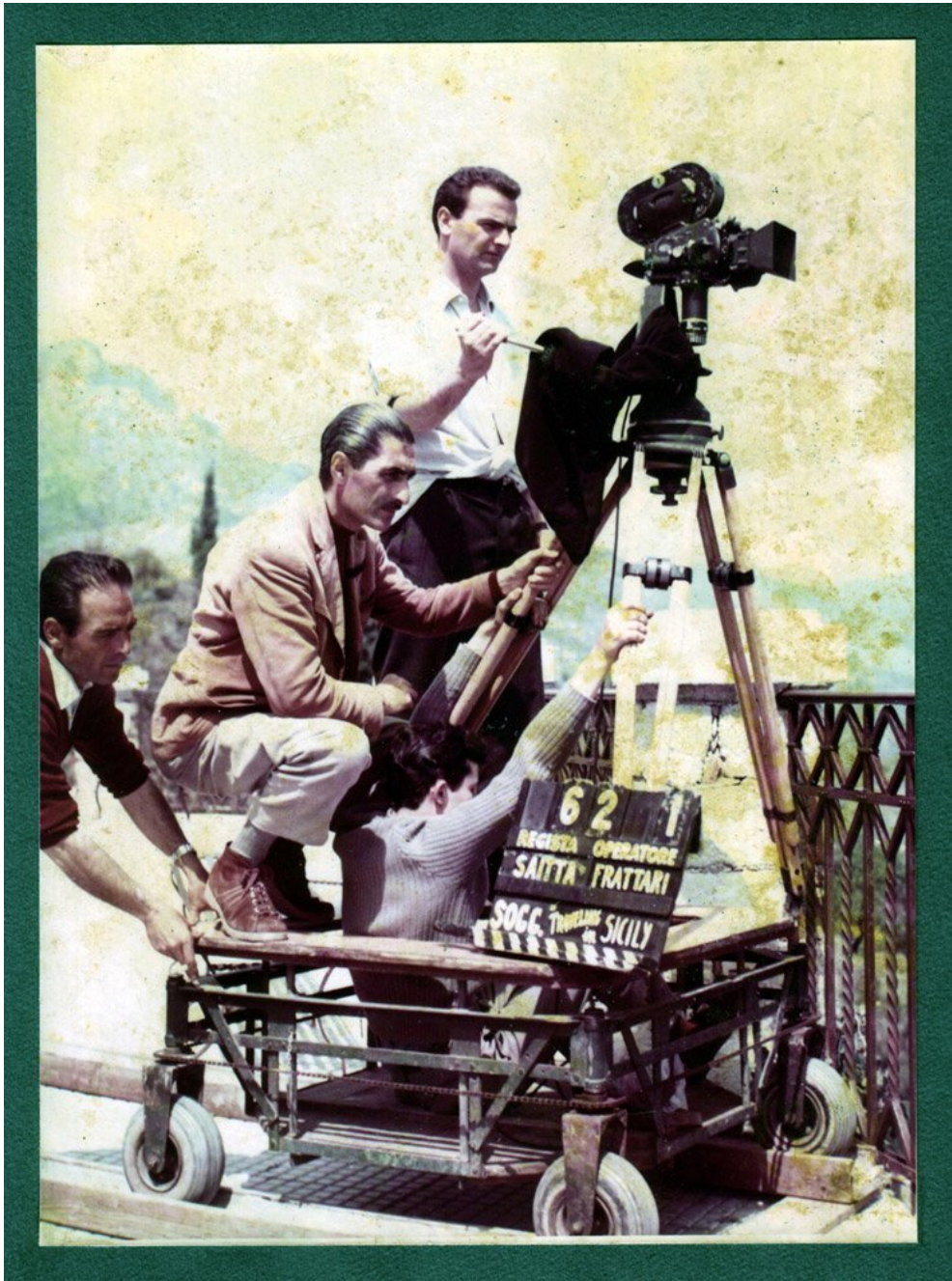
Produzione, Regia, Fotografia Ugo Saitta
Direzione di produzione Rita Consoli

4) **SOTTO IL SOLE DI ROMA**

(*ITALIA*)

Produzione Sandro Ghenzi (UNVERSALCINE)
Regia Ermanno Donati
Soggetto Nino Rota
Fotografia Renato Castellani, Sergio Amidei
Canzoni e musica Renato Castellani, Fausto Tozzi
Direttore di produzione Renato Castellani

Figura 17.67: Il programma della Biennale di Venezia del 1948

6 2 1
REGISTA OPERATORE
SAITTA FRATTARI
SOGG. TRAVELLING IN SICILY

17.4.2 Le foto di scena

L'attore Tuccio Musumeci, grande amico e collaboratore di Saitta, ha messo a disposizione un prezioso album fotografico con le foto di scena degli episodi di *Volto di Sicilia* e *Lo voglio maschio.*

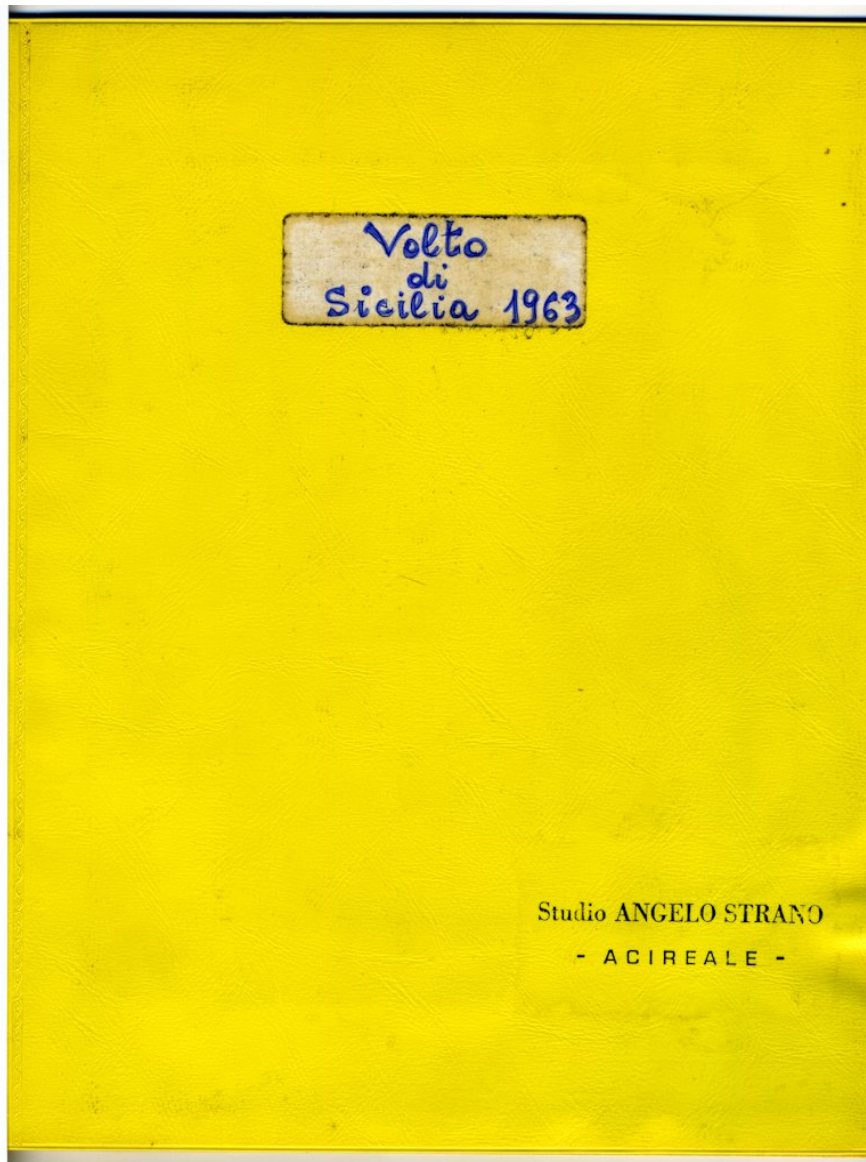



Figura 17.68: La copertina dell'album fotografico

Figura 17.69: *Volto di Sicilia n. 2*

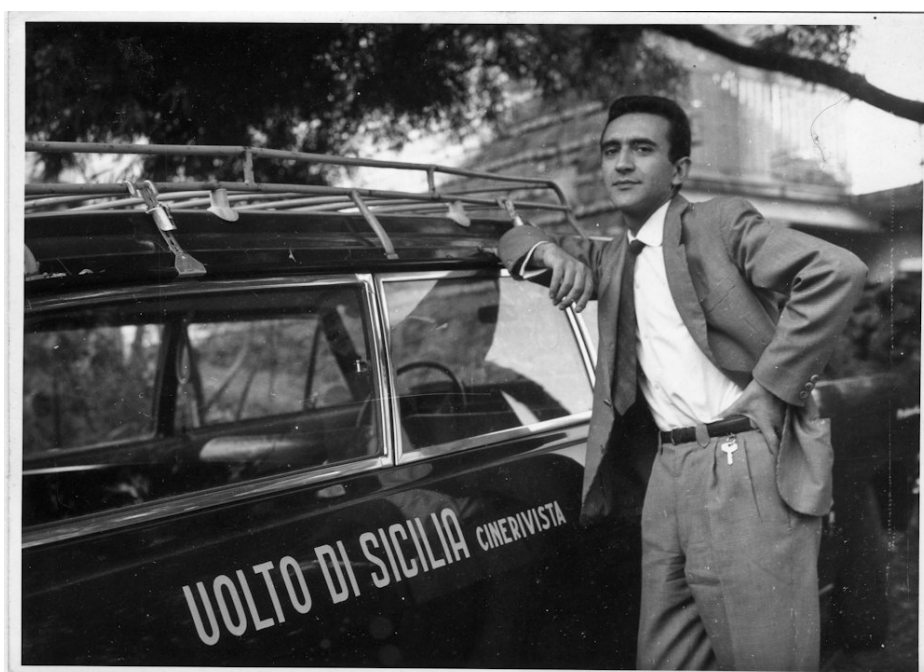



Figura 17.70: Tuccio Musumeci e l'auto di servizio di *Volto di Sicilia*

Figura 17.71: Tuccio Musumeci in *Volto di Sicilia n. 3,* episodio *Uno contro tutti e viceversa*

17.4.3 Archiviazione

La prima archiviazione dei files master ad altissima risoluzione è stata effettuata su un Hard Disk di rete della capienza di 4 terabyte. Successivamente, le immagini sono state ritagliate ed elaborate e restaurate digitalmente. Insieme alle immagini ad alta risoluzione, è stata prodotta una copia a bassa risoluzione

consultabile a schermo che è stata parzialmente messa in rete sul sito www.siciliadocumentaria.it, di cui si parlerà più avanti.

Gli hard disk FireWire utilizzati per l'archiviazione in triplice copia sono ubicati in differenti locazioni per garantire una conservazione sicura.

17.5 Schede tecniche dei film di Ugo Saitta

Attraverso lo studio dei materiali raccolti, è stato possibile ricostruire la filmografia completa di Ugo Saitta. È stata compilata una scheda tecnica completa delle opere rintracciate come disponibili presso la Filmoteca regionale siciliana. Si tratta di un elenco che indica i dati essenziali di ogni film come la durata, l'anno di produzione e la *crew* tecnica. La grande ricchezza di titoli e di temi affrontati, permette di delineare un percorso variegato e di grande curiosità anche per il pubblico dei non addetti ai lavori, di seguito elencato.

Teste di legno (1937, 11')
direzione della produzione: Rita Consoli
soggetto e sceneggiatura: Carmelo D'Arrigo
scenografia e costumi: Carolina Simeoli
animatore dei pupazzi: Ugo Saitta
operatore: Giuseppe Consoli

segretaria di produzione: Bianca Fichera

musica: Pier Capponi

regìa: Ugo Saitta

Zolfara (1947, 11')

Herald Pictures

direzione della produzione: Rita Consoli

commento parlato: Giampiero Pucci

elaborazione musicale: Virgilio Chiti

produzione, regìa e fotografia: Ugo Saitta

presentato alla IX Mostra Internazionale d'Arte Cinematografica di Venezia

L'Etna è bianco (1947, 10')

X Film Industria Cinematografica Siciliana, Catania

direzione della produzione: Rita Consoli

testo e commento parlato: Giampiero Pucci

commento musicale: Virgilio Chiti

voce: Arnoldo Foà

soggetto, sceneggiatura, fotografia e regìa: Ugo Saitta

registrazione sonora: cineluce

Sant'Agata (1949, 10')

direzione della produzione: Rita Consoli

commento parlato: Giampiero Pucci
elaborazione musicale: Virgilio Chiti
produzione, fotografia e regìa: Ugo Saitta

La valle dei templi (1950, 8')
direzione della produzione: Rita Consoli
commento parlato: Giampiero Pucci
elaborazione musicale: Virgilio Chiti
direzione della fotografia: Ugo Saitta
operatori: Tonino Bucci, Giuseppe Consoli
produzione e regìa: Ugo Saitta
ripresa diretta sonora: tarantella *Pi li cummari* di Franco Li Causi, eseguita dai gruppi folkloristici della Sicilia

Mito e realtà di Siracusa (1951, 9')
direzione della produzione: Rita Consoli
commento parlato: Giampiero Pucci
elaborazione musicale: Virgilio Chiti
direzione della fotografia: Ugo Saitta
operatori: Tonino Bucci, Giuseppe Consoli
produzione e regìa: Ugo Saitta

Sciara (1953, 10')
direzione della produzione: Rita Consoli

commento parlato: Giampiero Pucci
elaborazione musicale: Virgilio Chiti
operatore: Carmelo D'Arrigo
produzione e regìa: Ugo Saitta

Il carretto siciliano (1954, 7'30")
direzione della produzione: Rita Consoli
operatore: Giuseppe Consoli
commento parlato: Giuseppe Berretta
regìa: Ugo Saitta

I pupi siciliani (1955, 8')
soggetto e sceneggiatura: Ugo Saitta e Giuseppe Berretta
ricreazioni pittoriche: Sebastiano Milluzzo
direzione della produzione: Rita Consoli
operatore: Giuseppe Consoli
tecnico del suono: Enrico Di Grado
produzione e regìa: Ugo Saitta
con la partecipazione dell'attore Giuseppe Castelli
interni realizzati in ripresa sonora diretta nel teatro dei pupi "Etna" di Catania

La barca siciliana (1958, 9')
soggetto e sceneggiatura: Ugo Saitta

direzione della produzione: Rita Consoli

operatore: Giuseppe Consoli

aiuto regìa: Umberto Ecora

adattamento musicale: Pietro Marletta

testo del commento: Giuseppe Beretta

regìa: Ugo Saitta

La riviera dei tre golfi (1960, 9')

direzione della produzione: Rita Consoli

operatore: Tonino Ferrara

commento: Giampiero Pucci

commento musicale: Pietro Marletta

regìa: Ugo Saitta

Traveling in Sicily (1960, 10')

soggetto e sceneggiatura: Ugo Saitta e Giuseppe Berretta

direzione della produzione: Rita Consoli

musiche originali: Franco Cristina

operatore: Benito Frattari

con Alberto Testa e Loredana Furno

regìa: Ugo Saitta

Itinerario Etna (1960, 10')

Cineproduzione Ugo Saitta

testo del commento parlato: Giampiero Pucci
commento musicale: Pietro Marletta
fotografia: Ugo Saitta, Benito Frattari
regìa: Ugo Saitta

Siace 67 (1967, 8')
Cineproduzione Ugo Saitta
direzione della produzione: Rita Consoli
commento parlato: Gaetano Caponetto
voce: Riccardo Paladini
elaborazione musicale: Pietro Marletta
fotografia e regìa: Ugo Saitta

Etna quota 3000 (1969, 8')
Cineproduzione Ugo Saitta
organizzazione della produzione: Rita Consoli
direttore della fotografia: Ugo Saitta
operatore alla macchina: Angelo Strano
commento musicale: Mario Molino
regìa: Ugo Saitta

Storia su due ruote (1976, 11')
direttore della fotografia: Ugo Saitta
operatore alla macchina: Angelo Strano

commento musicale: Mario Molino
pittore di carretti: Domenico Di Mauro
cantastorie: Vito Santangelo

Civiltà contadina (1977, 11')
cineproduzione Ugo Saitta
organizzazione della produzione: Rita Consoli
direttore della fotografia: Ugo Saitta
operatore alla macchina: Angelo Strano
aiuto regista: Riccardo Manaò
testo del commento parlato: Gaetano Caponetto
commento musicale: Mario Molino
regìa: Ugo Saitta

La festa dei poveri (1978, 11')
Cineproduzione Ugo Saitta
organizzazione della produzione: Rita Consoli
direttore della fotografia: Ugo Saitta
operatore alla macchina: Angelo Strano
commento musicale: Mario Molino
regìa: Ugo Saitta

È Sicilia (1981, 11') incompleto
organizzazione della produzione: Rita Consoli

commento musicale: Mario Molino

regìa: Ugo Saitta

Schede tecniche ***Volto di Sicilia – cinerivista***

n° 0 (1951, 8')

direzione della produzione: Rita Consoli

commento parlato: Giampiero Pucci

elaborazione musicale: Virgilio Chiti

operatore: Carmelo D'Arrigo

produzione e regìa: Ugo Saitta

n°1 (1962, 10')

direttore: Ugo Saitta

redattore capo: Giuseppe Fava

redattori: Giuseppe Pappalardo e Gaetano Zappalà

direzione della produzione: Rita Consoli

operatore: Antonio Ferrara

commento musicale: Pietro Marletta

n°2 (1963, 9')

episodi: *Una nuova dimensione*, *Una ricchezza da difendere*, *Calcio follia*

direttore: Ugo Saitta

redattore capo: Giuseppe Fava

direzione della produzione: Rita Consoli

operatore: Antonio Ferrara

aiuto regista: Mario Ventimiglia

commento musicale: Pietro Marletta

n°3 (1963, 9')

episodi: *Un paradiso per Bacco*, *Uno contro tutti e viceversa*, *Gli aquiloni del lago*

direttore: Ugo Saitta

redattore capo: Giuseppe Fava

redattori: Gerardo Farkas, Orazio Ferrara

direzione della produzione: Rita Consoli

operatore: Antonio Ferrara

scenografa: Hilda Agnini

commento musicale: Pietro Marletta

sketch di: Tuccio Musumeci, Carmela Trovato, Riccardo Mangano, Turi Scalia,

Mario Piazza

n°5 (1965, 9')

episodi: *Il sole a spicchi*, *Una spiaggia internazionale*

direttore: Ugo Saitta

redattore capo: Giuseppe Fava

direzione della produzione: Rita Consoli

n°6 (1967, 8')

episodi: *Pirandello*, *Prospettive nuove per un'arte antica*, *Alla ricerca del mare perduto*

direttore: Ugo Saitta

redattore capo: Giuseppe Fava

direzione della produzione: Rita Consoli

n°7 (1968, 8')

episodi: *Il terremoto dei poveri*, *Dall'albero alla carta*, *L'arcipelago del sole*

direttore: Ugo Saitta

redattore: Gaetano Caponetto

organizzatrice: Rita Consoli

elaboratore musicale: Pietro Marletta

speaker: Riccardo Paladini

n°8 (1968, 8')

episodi: *I frutti del sole*, *Sicilia chiama mondo*

direttore: Ugo Saitta

redattore: Gaetano Caponetto

organizzatrice: Rita Consoli

elaboratore musicale: Pietro Marletta

speaker: Riccardo Paladini

n°9 (1969, 8')

episodi: *Sicilia: un ponte con il terzo mondo*, *Vacanze nel silenzio*

direttore: Ugo Saitta

redattore: Gaetano Caponetto

organizzatrice: Rita Consoli

elaboratore musicale: Pietro Marletta

speaker: Riccardo Paladini

17.6 Il documentario ritrovato: *La terra di Giovanni Verga*

La terra di Giovanni Verga è un'opera del 1953 sconosciuta alla stessa figlia del regista catanese e al responsabile del fondo Saitta, presso la Filmoteca Regionale di Palermo, il dott. Marcello Alajmo. Inizialmente si pensava fosse un documentario con un doppio titolo. Come spesso capita per le opere con una distribuzione accidentata, esistono infatti diverse copie di un unico film che registrano titoli alternativi. Effettuando però una ricerca più accurata, a livello nazionale, grazie alla storica del cinema dott.ssa Annamaria Licciardello, è stata scoperta l'esistenza del negativo di questo "documentario fantasma", conservato presso la Cineteca del Centro Sperimentale di Cinematografia di Roma. Dopo un primo controllo del negativo, che è in un discreto stato di conservazione, ma è un "infiammabile" e quindi per legge non può essere maneggiato né trasportato, si è proceduto alla richiesta della sua digitalizzazione, stampandone direttamente dal negativo una copia video in formato *Digital Betacam* (in alta definizione, per la sua conservazione in digitale) e una copia dvd, per la sua diffusione in rassegne e per motivi di studio. L'operazione è stata finanziata da Fuoricircuito e dalla stessa Cineteca Nazionale e patrocinata dalla Facoltà di Lettere e Filosofia.

17.7 Un documentario sui documentari di Saitta: *Con la Sicilia negli occhi*

Tutti i materiali recuperati, restaurati e catalogati hanno permesso la realizzazione del documentario *Con la Sicilia negli occhi. Il cinema documentario di Ugo Saitta*. Prodotto da Fuoricircuito e realizzato con la partecipazione del laboratorio multimediale di sperimentazione audiovisiva della Facoltà di Lettere e Filosofia dell'Università degli Studi di Catania (la.mu.s.a.), è stato finanziato dall'Assessorato al Turismo, Sport e Spettacolo della Regione Sicilia, con la collaborazione della Sicilia Film Commission, grazie al Fondo Regionale per il Cinema e l'Audiovisivo del 2008. Per la produzione del documentario ci si è valsi anche della preziosa collaborazione della Filmoteca Regionale Siciliana, che detiene e amministra il patrimonio cinematografico di Ugo Saitta, e della Cineteca Nazionale di Roma.

Con la Sicilia negli occhi contiene interventi di coloro che hanno conosciuto Ugo Saitta, che con lui hanno lavorato e collaborato o che meglio contribuiscono a definire il valore della sua produzione: Riccardo Manaò, Gabriella Saitta, Tuccio Musumeci, Leo Gullotta, Sebastiano Gesù, Adriano Aprà, Marco Bertozzi, Enrico Magrelli, Giuseppe Giarrizzo, Ivano Lollo e Mario Musumeci.

La crew tecnica che ha lavorato al documentario è composta da Alessandro De Filippo (fotografia e regìa), Ivano Mistretta (produzione), Carmelo Emmi (montaggio), Carlo Natoli (musiche

originali), Giuseppe Giudice, Giovanni Tomaselli (riprese), Alessandro De Caro (suono), Stefania Santagati (suono in presa diretta), Roberta Caruso, Chiara D'Amico (edizione), Stephen Conway (edizione dialoghi), Roberta Incatasciato (assistente al montaggio), Monica Saso (grafica e compositing), Antonio Caia (foto di scena), Felicita Platania (assistente di produzione), Antonio La Cognata (ispezione di produzione), Maria Vittoria Rasà (archivi), Valentina Sciuto (segreteria organizzativa).

Figura 17.72: Gabriella Saitta sul set

Figura 17.73Riccardo Manaò sul set

17.8 Un progetto sul documentario siciliano in rete: siciliadocumentaria.it

Data la quantità di documenti raccolti e la necessità di condividere i risultati di questo progetto di ricerca sul cinema di Ugo Saitta, è stato creato il sito siciliadocumentaria.it. Questo progetto vuole diventare un riferimento importante sulla rete, un vero e proprio archivio di documenti e informazioni sui documentari sulla Sicilia e sui documentaristi.

Figura 17.74 La homepage di siciliadocumentaria.it

Figura 17.75 I filmati di Ugo Saitta su siciliadocumentaria.it

18. Conclusioni

Gli studi sulla digitalizzazione e la conservazione delle fonti digitali dimostrano come, sempre più, l'integrazione tra la storia e la tecnologia e l'informatica muovano il lavoro dello storico verso una direzione nuova. È un processo in corso e, probabilmente, non sarà facile vederne una conclusione, considerato il fatto che l'innovazione tecnologica ha una velocità di crescita molto rapida.

I progetti descritti, che sono il risultato di una vera e propria ricerca sulle modalità di conservazione, dimostrano come anche chi si occupi di ricerca o metodologia della ricerca storica debba inventare nuovi modi di presentazione dei contenuti. La creazione del digital repository per i manoscritti di Federico De Roberto basato su grid, la digitalizzazione che permette la libera consultazione di un testo raro come il *Vojage pittoresque* di Houel, la creazione di prodotti multimediali come *I cantieri di De Carlo a Catania*, lo studio del video di Ugo Saitta e la produzione di un documentario che permette di conoscere i suoi modi d'indagine sul reale sono tutti esempi di lavori condotti sul versante tecnico e di significativa integrazione tra tecnologie scientifiche e informatiche, multimedialità e conoscenza umanistica.

Ogni progetto di digitalizzazione o multimediale va condotto rispettando accuratamente gli attuali standard e le linee guida internazionali, dando vita a strumenti informatici per l'umanistica che possano essere accolti dalla comunità di ricerca coniugando perfettamente le due parole chiave che sono alla base del trattamento digitale del patrimonio culturale: conservazione e accesso. La salvaguardia dei patrimonio è garantita dalla digitalizzazione dei materiali, mentre l'accesso ai documenti da parte della comunità umanistica è possibile grazie alla rete e alla possibilità di condivisione di qualunque tipo di mezzo di comunicazione.

La ricerca storica, le scienze umanistiche, dotandosi di strumenti estremamente avanzati in grado di potenziare i propri ambiti di ricerca, si aprono a collaborazioni interdisciplinari che, possono produrre risultati sorprendenti.

Nuovi modi di comunicare cultura umanistica e di produrre conoscenza sono oggi possibili attraverso le risorse informatiche e multimediali. Dalla digitalizzazione del patrimonio culturale alla conservazione e alla diffusione possono davvero aggiungere valore alla loro già smisurata ricchezza, investendo sulla computabilità umanistica che potrebbe condurre con successo, in un futuro per niente lontano, verso una vera e propria rivoluzione: quella della *humanities computer science*.

Bibliografia

AA.VV., *Dublin Core Metadata Element Set* (versione 1.1), trad. it. dell'Istituto Centrale per il Catalogo Unico, ICCU, <http://www.iccu.sbn.it/genera.jsp?id=116>

AA.VV., *EAD. Descrizione archivistica codificata. Dizionario dei marcatori*, a cura dell'Istituto Centrale per il Catalogo Unico, ICCU, Roma 2005

AA.VV., *Futuro delle memorie digitali e patrimonio culturale. Atti del convegno internazionale.* Firenze 16-17 ottobre 2003, a cura dell'Istituto Centrale per il Catalogo Unico, ICCU, Roma 2004

AA.VV., *Linee guida per la digitalizzazione del materiale cartografico*, a cura del Gruppo di lavoro sulla digitalizzazione del materiale cartografico dell'Istituto Centrale per il Catalogo Unico, ICCU, Roma 2006

AA.VV., *Linee guida per per la digitalizzazione del materiale fotografico*, a cura del Gruppo di lavoro per la catalogazione dei documenti fotografici dell'Istituto Centrale per il Catalogo Unico, ICCU, Roma 2006

AA.VV., *Linee guida per la digitalizzazione di bandi, manifesti e fogli volanti,* a cura del Gruppo di lavoro per la digitalizzazione di bandi, manifesti e fogli volanti dell'Istituto Centrale per il Catalogo Unico, ICCU, Roma 2006

AA.VV., *Memorie digitali. Rischi ed emergenze*, a cura dell'Istituto Centrale per il Catalogo Unico, ICCU, Roma 2005

AA.VV., Istituto Centrale per il Catalogo e la Documentazione (ICCD), *Normativa per l'acquisizione digitale delle immagini fotografiche*, 1998 <http://www.iccd.beniculturali.it/Catalogazione/standard-catalografici/aquisizione-digitale-delle-immagini-fotografiche>

ABBATTISTA G., ZORZI A. (a cura di), *Il Documento immateriale. Ricerca storica e nuovi linguaggi*, in «L'Indice del mese», n.4, maggio 2000, disponibile online su <http://lastoria.unipv.it/dossier/index.html>

BERARDI E., MARSICOLA C., *L'utilizzo della digitalizzazione nell'ICCD: stato dell'arte e prospettive*, «Digitalia», Numero 0, Anno I, 2005, pagg. 97-100

BIANCONI F., *Segni digitali. Sull'interpretazione e il significato della tecnologia digitale per la conservazione dei beni culturali*, Morlacchi, Perugia 2005

BUZZETTI D., *L'Informatica Umanistica come disciplina teorica*, in «Griseldaonline», <http://www.griseldaonline.it/informatica/buzzetti.htm>

CAFFO R., *La digitalizzazione del patrimonio culturale in Europa: il piano di lavoro congiunto della presidenza greca e della presidenza italiana dell'unione europea*, in «AIB notizie», anno XV, n. 11, 2003, pagg. 16-17

CALANCA D., *Percorsi di storiografia digitale,* in «Storia e Futuro. Rivista di Storia e Storiografia», n.6, Maggio 2005

CANALI D., *Standard per la biblioteca digitale. Nuovi linguaggi di codifica per l'informazione bibliografica*, Editrice Bibliografica, Milano 2006

CAPALDI D., ILARDI E., RAGONE G., *Comunicare la memoria. Le istituzioni culturali europee e la rete*, Liguori, Napoli 2008

CIOTTI F., *L'informatica umanistica in Italia. Luci e ombre*, in «Griseldaonline», <http://www.griseldaonline.it/informatica/ciotti.htm>

COMITATO MAG, *MAG: Metadati Amministrativi e Gestionali. Manuale Utente* (versione 2.0.1), a cura di E. Pierazzo, ICCU, <http://www.iccu.sbn.it/upload/documenti/Manuale.pdf>

COMMISSIONE EUROPEA, *Lund Principles*, Svezia, 2001, <http://cordis.europa.eu/ist/digicult/lund-principles.htm>, versione italiana disponibile su <ftp://ftp.cordis.lu/pub/ist/docs/digicult/lund_principles-it.pdf>

CRISCIONE A., NOIRET S., SPAGNOLO C., VITALI S., *La storia a(l) tempo di Internet*, Patron, Bologna 2004

DE FRANCESCO G., *Linee guida tecniche per i programmi di creazione di contenuti culturali digitali* (versione italiana 1.8), a cura del Gruppo di lavoro italiano *Interoperabilità e servizi* del progetto Minerva <http://www.minervaeurope.org/structure/workinggroups/servprov/documents/technicalguidelinesita1_8.pdf>

FALCHETTA P., *Guida breve alla digitalizzazione in biblioteca*, in «Biblioteche oggi», n. 9, anno VIII, novembre 2000, pagg. 52-67.

FERRI P., *La rivoluzione digitale*, Mimesis, Milano 1999

FIORMONTE D., *Informatica Umanistica: rappresentanza, statuto teorico e rifondazione*, in «Griseldaonline», <http://www.griseldaonline.it/informatica/fiormonte_risposta.htm>.

GENOVESE N., GESÙ S., *E venne il cinematografo. Le origini del cinema in Sicilia*, Maimone, Catania 1995

GESÙ S., *La Sicilia della memoria. Cento anni di cinema documentario nell'isola,* Maimone, Catania 1999

GESÙ S., *L'Etna nel cinema. Un vulcano di celluloide,* Maimone, Catania 2005

GORI E., *La storiografia digitale in Italia - Orizzonti e prospettive* in «Historied.net - studi e risorse per la storia dell'educazione», n.1, marzo 2007

KUHN T., *La struttura delle rivoluzioni scientifiche*, Einaudi, Torino 1999

LOTTI F., *Le qualità delle immagini nei progetti di digitalizzazione*, in «Digitalia», n.2, Anno II, 2006, pagg. 22-37.

MAGLIANO C., *Lo standard nazionale dei metadati gestionali amministrativi*, in «Digitalia», n.0, Anno I, 2005, pagg. 34-46.

MALDONADO T., *Memoria e conoscenza. Sulle sorti del sapere nella prospettiva digitale*, Feltrinelli, Milano 2005

NOIRET S. (a cura di), *Linguaggi e siti: La Storia On line*, in «Memoria e ricerca. Rivista di storia contemporanea», fascicolo monografico, a.VII, n.s, n.3, Carocci, Roma 1999

ORLANDI T., *Proposta: Informatica applicata alle discipline umanistiche (ovvero: Informatica umanistica)*, in «Griseldaonline», <http://www.griseldaonline.it/informatica/orlandi.htm>.

NEGROPONTE N., *Essere digitali*, Sperling & Kupfer, Milano 2004

ORTOLEVA P., *Mediastoria. Comunicazione e cambiamento sociale nel mondo contemporaneo*, Il Saggiatore, Milano 2002

ORTOLEVA P., *Il secolo dei media. Riti, abitudini, mitologie*, Il Saggiatore, Milano 2009

ORTOLEVA P., *Testimone infallibile, macchina dei sogni: il film e il programma televisivo come fonte storica*, in «Passato e presente. Rivista di storia contemporanea», luglio-dicembre 1984

PARISINI R., *Indagini di storiografia digitale*, in «Storia e Futuro. Rivista di Storia e Storiografia», n.6, Maggio 2005

PICCININNO M., *Parma: Conferenza europea di Minerva: i rappresentanti europei per la digitalizzazione a confronto*, in «AIB Notizie», anno XV, n. 11, 2003, pagg. 11-12

RAGAZZINI D. (a cura di), *La storiografia digitale*, Utet, Torino 2004

RONCAGLIA G., *I progetti internazionali di digitalizzazione bibliotecaria: un panorama in evoluzione*, in «Digitalia», n.1, Anno II, 2006, pagg. 11-30.

RONCAGLIA G., *Informatica umanistica: le ragioni di una disciplina*, in «Griseldaonline», <http://www.griseldaonline.it/informatica/roncaglia_secondo.htm>

SALARELLI A., TAMMARO A.M., *La biblioteca digitale*, Editrice Bibliografica, Milano 2006

TAMMARO A.M., *Che cos'è una biblioteca digitale?*, in «Digitalia», Numero 0, Anno I, 2005, pagg. 14-33

TOMASI F., *Studia humanitatis e nuove tecnologie. La scienza dell'informazione come presupposto metodologico*, in «Griseldaonline», 12/11/2006, <http://www.griseldaonline.it/informatica/3tommasi.htm>

VITALI S., *Passato digitale. Le fonti dello storico nell'era del computer*, Mondadori, Milano 2004

Sitografia

http://www.iccu.sbn.it
Sito dell'Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU)

http://www.aib.it
Sito dell'Associazione Italiana Biblioteche

http://www.bibliotecaitaliana.it
Biblioteca digitale di testi rappresentativi della tradizione culturale e letteraria italiana dal Medioevo al Novecento, promossa dal Centro interuniversitario Biblioteca italiana telematica (CiBit), e gestita presso l'Università di Roma "La Sapienza"

http://books.google.it
Sito di Google Books

http://www.culturadigitale.it
Portale con articoli e risorse di cultura digitale e per le biblioteche digitali

http://www.digitalpreservationeurope.eu
Sito del progetto Digital Preservation Europe

http://www.digitalia.sbn.it

Rivista online sul digitale nei beni culturali, edita dall'ICCU

http://www.griseldaonline.it

Portale di letteratura dedicato all'informatica umanistica

http://www.interpares.org/

Sito del progetto InterPARES

http://www.minervaeurope.org

Sito del progetto europeo Minerva

http://www.michael-culture.org/it

Versione italiana del sito del progetto Michael.

http://www.w3.org/

Sito del *World Wide Web Consortium*

Glossario

A

A. Formato di una tipologia di documenti definita dallo standard →*ISO216*.

A/D converter. Dispositivo che converte i dati dal segnale →*digitale* a quello →*analogico*.

Acquisizione. Procedura volte alla ricezione di dati, provenienti da fonte esterna, nel sistema in uso.

Algoritmo. Metodo per la soluzione di un problema adatto a essere implementato sotto forma di programma.

Alpha channel. Canale aggiuntivo (non obbligatorio e non sempre presente), che descrive il grado di trasparenza/opacità di ogni determinato →*pixel*, con un valore numerico variabile a seconda dei formati di output.

Analogico. Riferito a un segnale, indica che i valori utili che lo rappresentano sono continui, non discreti.

Anodo. Elettrodo sul quale avviene una reazione di ossidazione.

API. Acronimo dell'inglese *Application Programming Interface* (interfaccia di programmazione di un'applicazione), indica ogni insieme di procedure disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per un determinato compito.

ASCII. Acronimo di *American Standard Code for Information*

Interchange (Codice Standard Americano per lo Scambio di Informazioni). È un sistema di codifica dei caratteri a 7 →*bit*.

B

Banca dati. Vedi →*database*.

Back end. Stadio finale di un processo.

Batch. insieme di comandi o programmi, tipicamente non interattivi, aggregati per l'esecuzione.

Bit. Dall'inglese *binary unit*, unità di misura dell'informazione, definita come la quantità minima che serve a discernere tra due possibili alternative equiprobabili.

BMP. Formato dati utilizzato per la rappresentazione di immagini raster sui sistemi operativi Microsoft Windows.

Born digital. Documento creato ed esistente soltanto in digitale.

BPP. Acronimo di *bit per pixel*, unità di misura della →*profondità di colore*.

Brightness. Quantità totale di luce che una sorgente luminosa appare emettere o che appare riflessa da una superficie.

Buffer. Letteralmente "tampone", zona di memoria usata temporaneamente per l'input o l'output dei dati oppure per velocizzare l'esecuzione di alcune operazioni.

Bus. Canale che permette a periferiche e componenti del sistema di dialogare tra loro.

Byte. Sequenza composta da 8 →*bit*, in grado di rappresentare 256 possibili valori.

C

C++. Linguaggio di programmazione con caratteristiche orientate all'oggetto, comunemente utilizzato per lo sviluppo di sistemi e software.

Calibrazione. Operazione attraverso cui uno strumento viene regolato in modo da migliorarne l'accuratezza. L'operazione richiede il confronto con delle misure di riferimento prodotte utilizzando uno strumento campione.

Campionamento. Tecnica che consiste nel convertire un segnale continuo nel tempo in un segnale discreto, valutandone l'ampiezza a intervalli di tempo regolari. In questo modo, a seguito di una successiva operazione di quantizzazione e conversione, è possibile ottenere una stringa digitale (discreta nel tempo e nell'ampiezza) che si approssimi quella continua e originaria.

Canale. Porzione dello spettro elettromagnetico, all'interno della quale è possibile trasmettere il segnale generato dalla sorgente di informazione che vi accede.

Catodo. Elettrodo sul quale avviene una reazione di riduzione.

CCD. Acronimo di *Charge Coupled Device* , è un circuito integrato formato da una riga, o da una griglia, di elementi semiconduttori (*photosite*) in grado di accumulare una carica elettrica (*charge*), proporzionale all'intensità della radiazione elettromagnetica che li colpisce.

CD-ROM acronimo di *Compact Disc Read Only Memory (→ROM)*, è un supporto in policarbonato di 12 centimetri, usato per la memorizzazione di dati generici. Si ottengono per stampa, con appositi macchinari industriali. I CD-ROM di più larga distribuzione hanno una capienza di 74 minuti/650 MB e di 80 minuti/700 MB, mentre più rari sono i formati da 90 minuti/800

MB e da 100 minuti/870 MB.

CIE. →*Spazio colore,* definito nel 1931 dalla *Commission Internationale de l'Eclairage*, che comprende tutte le tinte visibili dall'occhio Umano. Lo spazio così definito diviene tridimensionale e rappresentato mediante coordinate XYZ, basate sull'utilizzo di tre colori primari che, miscelati tra loro in sintesi additiva, permettono di ottenere tutti i colori che l'occhio umano può percepire.

Client. Computer collegato a un →*server* tramite una rete informatica (locale o geografica), al quale richiede uno o più servizi, utilizzando protocolli di rete.

Client side. Letteralmente, "dal lato del cliente". È detto di un'operazione che non è eseguita dal sistema remoto a cui ci si collega, ma dal sistema locale: l'utente riceve i dati a partire dai quali il proprio calcolatore ricaverà il risultato dell'operazione, che gli sarà infine mostrato. Il sistema complementare è →*server side.*

Cluster. In informatica, insieme di computer connessi tramite una rete telematica, allo scopo di distribuire una elaborazione molto complessa tra i vari computer che compongono il cluster. Un problema che richiede molte elaborazioni per essere risolto viene scomposto in sottoproblemi separati i quali vengono risolti in parallelo. Questo ovviamente aumenta la potenza di calcolo del sistema.

CMYK. Acronimo per *Cyan, Magenta, Yellow, BlacK*, modello di colore a quattro componenti, detto anche di quadricromia.

CMOS. Tipo di tecnologia utilizzata in elettronica per la progettazione di componenti digitali utilizzando transistor.

Codice binario. notazione che utilizza simboli binari o →*bit.*

Core. Nucleo elaborativo di un →*microprocessore.*

Color depht. Vedi →*profondità di colore.*

Colorchecker. Tabella di test per il colore, concepita secondo canoni scientificamente accurati, per determinate il corretto bilanciamento cromatico dei dispositivi di acquisizione, sia video che fotografici.

Compressione. Tecnica preposta alla riduzione del numero di →*bit* necessari per immagazzinare un'informazione, generalmente applicata a ridurre le dimensioni di un file. Può essere di tipo →*lossless* o →*lossy.*

Contrasto. Rapporto o differenza tra il valore più alto (punto più luminoso) e il valore più basso (punto più scuro) in un'immagine.

Copyleft. Licenza libera sul diritto d'autore.

Copyright. Forma di garanzia del diritto d'autore.

Cpu. Acronimo dell'inglese *Central Processing Unit* (unità centrale di elaborazione), denominata in genere nella sua implementazione fisica *processore.*

Crop. Operazione di ritaglio su un'immagine.

CRT. Acronimo del termine inglese *Cathode Ray Tube*, indica la tecnologia comunemente usata per la visualizzazione nei monitor e nei televisori, che consiste nel convogliare i raggi catodici su di una superficie sensibile, che ricostruisce l'immagine visibile. La misura dei monitor CRT si effettua sulla diagonale dell'area visibile, in pollici.

D

Dat. Acronimo del termine inglese *Digital Audio Tape*, è un mezzo per la registrazione e riproduzione di audio introdotto da Sony nel 1987.
Il DAT converte il suono ad una frequenza di campionamento più o meno elevata rispetto al CD (48, 44.1 o 32 kHz di frequenza, e 16 bit di quantificazione), senza compressione audio.

Database. Indica un archivio, strutturato in modo tale da consentire la gestione dei dati stessi (l'inserimento, la ricerca, la cancellazione e il loro aggiornamento) da parte di applicazioni software.

DDC. Acronimo dell'inglese *Display Data Channel***,** è uno standard VESA (*Video Electronics Standards Association*) che descrive una connessione fisica fra un display e una scheda video, al fine di poter permettere al display di poter trasmettere i suoi dati identificativi.

DDT. Acronimo dell'inglese *Document Type Definition* (definizione del tipo di documento), definisce le componenti ammesse nella costruzione di un documento →*XML*. Il termine non è utilizzato soltanto per i documenti XML ma anche per tutti i documenti derivati dall'→*SGML* tra cui il più noto è l'→*HTML*

Descreening. Procedimento attraverso il quale si ottiene, da una o più pellicole retinate, un file a tinta continua, cioè senza retino. Questo file può essere poi utilizzato con qualsiasi software di elaboraziamentoone immagini.

Diapositiva. Immagine fotografica positiva, in bianco e nero o a colori, su un supporto trasparente. Il suo nome deriva da *dia*, cioè attraverso, e *positiva*: infatti per poterla osservare occorre che la luce la attraversi.

Digitale. Termine che indica in informatica tutto ciò che viene rappresentato con numeri, o che opera manipolando numeri. Il termine deriva dall'inglese *digit*, che significa cifra.

Digitalizzazione. processo di conversione che, applicato alla misurazione di un fenomeno naturale o fisico, determina il passaggio dal campo dei valori continui a quello dei valori discreti, nei termini di un passaggio dall'→*analogico* al →*digitale*.

Disco rigido. Vedi →*hard disk*.

Disco ottico. Supporto di memorizzazione costituito da un disco piatto e sottile in genere di policarbonato trasparente. Al suo interno è inserito un sottile foglio metallico, in genere di alluminio, su cui vengono registrate e lette le informazioni tramite un raggio laser.

Display. Schermo video che rappresenta visivamente i dati forniti o elaborati da un'apparecchiatura elettronica.

Display data channel. Vedi →*DDC*.

Dithering removal. Vedi →*descreening*

Djvu, tecnologia sviluppata alla fine degli novanta da un gruppo di ricercatori dei laboratori della AT&T, con l'obiettivo di rendere possibile la realizzazione di raccolte →*digitali* di documenti ad alta qualità a colori.

Dorso digitale. Periferica contenente un sensore CCD/CMOS di grosse dimensioni ed elevata risoluzione, da utilizzare tramite adattatori su macchine fotografiche medio e grande formato.

Dorso a scansione. Usato in genere con le fotocamera di grande formato, funzionano come uno →*scanner piano*. Una singola riga di fotorecettori viene mossa da un motorino lungo il dorso, producendo come risultato finale un file di elevata risoluzione.

DMD. Acronimo dell'inglese *digital multilayer disk*, supporto ottico

sviluppato a partire dalla tecnologia usata per il →*FMD*.

Dpi. Acronimo dell'inglese *dots per inch*, indica la quantità di informazioni grafiche che possono essere rese da una stampante grafica o da uno schermo. Con dpi si esprime la quantità di punti stampati o visualizzati su una linea lunga un pollice (2,54 cm).

Driver. →*Software* che permette ad un sistema operativo di pilotare un dispositivo hardware.

Drum scanner. Scanner che utilizza sensori PMT (acronimo dell'inglese →*PhotoMultiplier Tubes*) costituiti da tre fotomoltiplicatori sensibili alle luci rossa, verde e blu (→*RGB*).

Dublin core. Sistema di →*metadati* costituito da un nucleo di elementi essenziali ai fini della descrizione di qualsiasi materiale digitale accessibile via rete informatica.

DVD-ROM, acronimo dell'inglese *Digital Versatile Disc Read Only Memory* (→*ROM)* ,è un supporto di memorizzazione di tipo ottico. Può avere uno o due lati scrivibili ed essere a singolo o doppio strato. Combinando queste caratteristiche può avere una capienza che va dai 4,7 GB del DVD-5 ai 18 GB del DVD-18.

E

Editing. processo attraverso cui si modificano immagini e altri media, con la correzione e altri tipi di modifiche.

End tag. Nei linguaggi di marcatura in genere, rappresenta la fine di un elemento. Un end tag inizia con "</" e finisce con ">".

eSATA, interfaccia standard di comunicazione seriale più performante rispetto a →*USB* e →*FireWire*. Può arrivare a 3Gb/sec (gigabit al secondo).

Esportazione. Processo di conversione di un file da un formato ad un altro.

F

File system. Insieme dei tipi di dati astratti necessari per la memorizzazione, l'organizzazione gerarchica, la manipolazione, la navigazione, l'accesso e la lettura dei dati.

FireWire. Interfaccia standard per un →*bus* seriale, che può supportare fino a 63 periferiche organizzate in una rete non ciclica. FireWire 400 specifica la velocità di trasferimento dati ai dispositivi di 100, 200, o 400 Mbps. FireWire 800 innalza la velocità della connessione a 800 Mbps.

Flatbed. Scanner piano capace di catturare foto, diapositive o negativi e tradurli in file digitali per l'elaborazione al computer.

FMD. Acronimo dell' inglese *Fluorescent Multilayer Disk*,formato di disco ottico che usa materiali fluorescenti piuttosto che riflettivi per archiviare i dati. L'uso della fluoresenza permette di raggiungere capacità di 1 terabyte, mantenendo sempre la taglia dei tradizionali supporti ottici.

Fotomoltiplicatore. Rivelatore elettronico di luce estremamente sensibile nell'ultravioletto, in luce visibile e nel vicino infrarosso. Il dispositivo è talmente sensibile da potere rilevare un singolo fotone.

Fotocamera. Apparecchio (comunemente chiamato macchina fotografica) usato per acquisire foto, singolarmente o in sequenza. Le fotocamere tradizionali catturano la luce su una pellicola fotografica o su una lastra fotografica. Le fotocamere digitali invece utilizzano l'elettronica, di solito un →*sensore* →*CCD* o →*CMOS* ,

per catturare le immagini, salvandole poi in apposite memorie.

Fotoritocco. Particolare forma di →*editing* che comprende l'insieme dei procedimenti che portano alla modifica di una fotografia.

Front-end. Nel campo della progettazione software, la parte di un sistema che gestisce l'interazione con l'utente o con sistemi esterni che producono dati di ingresso.

FTP. Acronimo di *File Transfer Protocol* (protocollo di trasferimento file), è un protocollo per la trasmissione di dati tra host basato su →*TCP*.

G

Gamma. Termine che in fotografia descrive la non linearità analogica della pellicola.

Gamma di densità. Inclinazione della curva sensitometrica, che è una proiezione della densità (ovvero il logaritmo dell'opacità) dell'immagine sulla pellicola rispetto al logaritmo dell'esposizione della pellicola alla luce.

GIF, acronimo per l'inglese Graphics Interchange Format, è un formato per immagini di tipo bitmap (→*BPM)* utilizzato nel World Wide Web (→*WWW)* , sia per immagini fisse che per le animazioni.

H

Hardware. Vocabolo creato a partire da due termini della lingua inglese, *hard* (duro) e *ware* (manufatto, oggetto). Definisce la parte fisica di un personal computer, ovvero tutte quelle parti magnetiche, ottiche, meccaniche ed elettroniche che gli consentono di funzionare.

Hard disk. Dispositivo utilizzato per la memorizzazione a lungo termine dei dati in un computer. È costituito da uno o più dischi in alluminio o vetro, rivestiti di materiale ferromagnetico in rapida rotazione e da due testine per ogni disco (una per lato), le quali, durante il funzionamento "volano" alla distanza di poche decine di nanometri dalla superficie del disco, leggendo e scrivendo i dati. La testina è tenuta sollevata dall'aria mossa dalla rotazione stessa dei dischi che può superare i 15.000 giri al minuto.

Hi-color. Abbreviazione per *High color*, modalità grafica (denominata spesso *migliaia di colori*), che memorizza le immagini nella →*ram* video, utilizzando due →*byte* per ogni →*pixel*. Con due byte, ossia 16 bit, il colore può essere rappresentato utilizzando tutti e 16 i bit (modalità più frequente) o solo 15 bit.

Html. Acronimo dell'inglese *Hyper Text Markup Language* (linguaggio di marcatura per ipertesti), è un linguaggio usato per descrivere i documenti ipertestuali disponibili nel →*World Wide Web*. L'HTML è un linguaggio di markup, ossia descrive il contenuto, testuale e non, di una pagina web.

Hue vedi →*tonalità*

I

ICC vedi →*International Color Consortium*

ICR. Acronimo dell'inglese *Intellligent Character Recognition*, indica il processo di riconoscimento intelligente dei caratteri per la conversione di un'immagine contenente testo in testo modificabile.

Importazione. procedimento di conversione dati nel formato nativo dell'applicazione utilizzata.

Indirizzo IP. Indirizzo numerico che identifica univocamente ogni computer collegato in rete (→*internet* o →*intranet*). Può essere statico (non cambia mai) oppure dinamico (cambia ad ogni connessione).

Information retrieval. Letteralmente *recupero delle informazioni*, è l'insieme delle tecniche utilizzate per il recupero mirato dell'informazione in formato elettronico.

Input. Indica in senso stretto il dato fornito ad una elaborazione e in senso più ampio l'insieme di elementi per formare qualcosa.

International Color Consortium. Organizzazione indipendente, fondata nel 1993, allo scopo di sviluppare e mantenere uno standard aperto, a livello di sistema operativo e multipiattaforma, per la gestione digitale del colore.

Internet. Prima e unica rete di computer mondiale ad accesso pubblico realizzata. Internet può essere vista come una rete logica di enorme complessità, appoggiata a strutture fisiche e collegamenti di vario tipo (fibre ottiche, cavi coassiali, collegamenti satellitari, doppino telefonico, collegamenti su radiofrequenza (WiFi), su ponti radio, su raggi laser e su onde convogliate su condotte elettriche o addirittura idrauliche) che interconnette un agente umano o automatico ad un altro agente tramite, praticamente, qualsiasi tipo di computer o elaboratore elettronico oggi o in futuro esistente o immaginabile.

Interpolazione. Procedimento per determinare in modo approssimato valori incogniti, partendo da valori noti.

Intranet. Rete locale (LAN), o raggruppamento di reti locali, usata all'interno di una organizzazione per facilitare la comunicazione e l'accesso all'informazione, che può essere ad accesso ristretto.

IP. Vedi →*indirizzo Ip*.

Ipertesto. Particolare tipo di testo informatico che permette la lettura a video grazie a un'interfaccia elettronica. Alcune parole chiave contenute appaiono marcate (sottolineate oppure evidenziate) e l'azione del clic con il mouse su di esse ha come conseguenza l'apertura di un altro documento correlato o lo spostamento su un'altra porzione dello stesso documento. Le parole chiave funzionano come collegamenti ipertestuali (hyperlink in inglese).

Iso 216, norma che descrive i formati di carta usati in numerosi paesi, specialmente in Europa.

J

Job. Unità di lavoro, in riferimento a compiti di vario tipo.

JPEG. Acronimo dell'inglese *Joint Photographic Experts Group* (comitato che ha definito il primo standard internazionale di compressione per immagini a tono continuo), definisce lo standard di compressione delle immagini fotografiche più utilizzato.

JPEG2000. Standard di compressione immagini basato sulla trasformazione →*wavelet* discreta.

L

LCD. Acronimo dell'inglese *Liquid Crystal Display*, è uno schermo sottile e leggero basato sulle proprietà ottiche di particolari sostanze denominate cristalli liquidi.

Login. Indica il processo di identificazione, genralmente attraverso l'inserimento del proprio nome utente e di una password.

Look up table. In image processing sono tabelle che collegano indici numerici (che vanno generalmente da 0 a 255) con valori di output.

Lossless. Classe di algoritmi di compressione dati che non porta a perdita di informazioni durante la fase di compressione/decompressione.

Lossy. Classe di algoritmi di compressione dati che porta a perdita di informazioni durante la fase di compressione/decompressione.

LUT. Vedi →*Look up table.*

LZW. Acronimo di *Lempel-Ziv-Welch*, i nomi dei suoi inventori, è un algoritmo che permette la compressione di dati senza perdita di informazioni (→*lossless*).

M

MAG. Metadati amministrativi gestionali.

Mainframe. Computer utilizzato per applicazioni critiche, per elaborare con alte prestazioni e alta affidabilità grandi moli di dati.

Mezzitoni. Tonalità in un'immagine che si trova al centro della gamma di toni, a metà del tono più chiaro e il tono più scuro.

Metadato. Letteralmente "dato su un (altro) dato", è l'informazione che descrive un insieme di dati.

Microprocessore. Implementazione fisica della →*CPU* in un singolo circuito integrato, utilizzata dalla quasi totalità dei computer e in altri dispositivi digitali.

Middleware. insieme di programmi informatici che fungono da intermediari tra diverse applicazioni, spesso utilizzati come supporto per applicazioni distribuite complesse.

Modello colore. Modello matematico astratto che permette di rappresentare i colori in forma numerica, tipicamente utilizzando tre o quattro valori o componenti cromatiche.

Moirè (effetto). Figura di interferenza, creata ad esempio da due griglie sovrapposta ad un certo angolo, o anche da griglie parallele con maglie distanziate in modo leggermente diverso.

Monitor. Apparecchio elettronico che consente la visualizzazione di immagini video statiche o in movimento.

Mbps. Acronimo per l'inglese *Megabit per second* (simbolo Mb/s o più impropriamente Mbit/s o Mbps), è un'unità di misura che indica la capacità (quindi velocità) di trasmissione dei dati su una rete informatica. Un megabit corrisponde a 1.000 kilobit, quindi a 1.000.000 bit.

N

Namespace. In italiano *spazio dei nomi*, collezione di nomi di entità, definite dal programmatore, omogeneamente usate in uno o più file sorgente. Lo scopo dei namespace è quello di evitare confusione ed equivoci nel caso siano necessarie molte entità con nomi simili. È presente esplicitamente nell' →*XML* e nel linguaggio di programmazione →*C++*.

Nuovi media. Mezzi di comunicazione di massa sviluppati posteriormente alla nascita dell'informatica e in correlazione ad essa.

O

OAIS. Acronimo per *Open Archival Information System*, è il nome di uno standard per l'archiviazione dei contenuti digitali. È stato creato dalla Nasa e permette a tutti coloro che lavorano a un progetto di visualizzare lo stesso archivio.

OCR. Acronimo dell'inglese *Optical Character Recognition* (riconoscimento ottico dei caratteri), è un sistema usato da programmi dedicati alla conversione di un'immagine contenente testo, solitamente acquisite tramite scanner, in testo digitale modificabile con un normale editor.

Offset. Tecnica di stampa basata sullo stesso principio della litografia. Anziché stampare il foglio a contatto diretto con la pietra o la lastra di alluminio microgranito, la stampa avviene attraverso l'impiego di tre cilindri a contatto tra loro.

On screen display. Letteralmente *display sullo schermo*, è un'immagine sovrapposta all'immagine sullo schermo, usata frequentemente nei moderni televisori, nei videoregistratori, lettori DVD per visulizzare informazioni come il volume, il canale e l'ora.

Open access. Letteralmente accesso aperto, è la libera disponibilità on-line di contenuti digitali. Esso viene principalmente utilizzato per gli articoli di riviste o di ricerche universitarie passate in rassegna da pari, pubblicate gratuitamente.

Open use. Licenza concessa dal detentore di un diritto d'autore, la cui caratteristica principale è generalmente la possibilità di ripubblicazione o riutilizzo dell'opera, poste alcune condizioni.

Open source. Letteralmente *sorgente aperto*, tipo di licenza usata per diversi tipi di software lasciati disponibili in codice-sorgente per ulteriori sviluppi. Spesso si intende, impropriamente, il software libero con questa dicitura.

Optical character recognition. Vedi →*OCR*.

Optoelettronico (gruppo). Sistema utilizzato negli scanner, composto da una lampada fredda che emette luce bianca per illuminare il foglio, creando dei fasci di luce che, raccolti da apposite lenti, vengono indirizzati al →*sensore*.

OSD. Vedi →*On screen display*.

Output. Dall'inglese *messo fuori*, indica in senso stretto il risultato di una elaborazione e in senso più ampio l'insieme dei risultati prodotti o semplicemente risultato.

P

Parser. →*Software* che, attraverso processi di parsing (o analisi sintattica), analizza uno →*stream* continuo in →*input* (letto per esempio da un file o una tastiera) in modo da determinare la sua struttura grammaticale grazie ad una data grammatica formale.

PCD. Acronimo dell'inglese *Photo CD format*, formato compresso sviluppato da Kodak e utilizzato per salvare immagini sui Kodak Photo CD.

PDF. Acronimo dell'inglese *Portable Document Format*, formato di file sviluppato da Adobe Systems nel 1993, per rappresentare documenti in modo indipendente dall'hardware e dal software utilizzati per generarli o per visualizzarli.

Periferica. Dispositivo →*Hardware* collegato al computer talvolta controllato da un proprio →*microprocessore*, le cui funzioni sono

controllate dal sistema operativo attraverso i relativi →*driver*.

Pellicola fotografica. Supporto costruito a strati, atto a conservare le immagini riprese a mezzo di una macchina fotografica analogica.

Photomultiplier tubes. Tubo fotomoltiplicatore (talvolta abbreviato in Pmt), rivelatore elettronico di luce estremamente sensibile nell'ultravioletto, in luce visibile e nel vicino infrarosso. Il dispositivo è talmente sensibile da potere rilevare un singolo fotone.

Piano (scanner). Dispositivo che impiega un gruppo →*optoelettronico* che scorre parallelamente al documento, mentre questo rimane immobile su un piano orizzontale.

Pixel. Contrazione della locuzione inglese *picture element*, elemento puntiforme che compone la rappresentazione di una immagine →*raster* nella memoria di un computer.

Profondità colore. Valore che indica la quantità di →*bit* necessari per rappresentare il colore di un singolo pixel in un'immagine →*raster*.

Proxy Computer che manda richieste per conto di altri. Un utente si può connettere ad un proxy e mandare richieste destinate ad altri server, il proxy le farà come se venissero da lui e rimanderà indietro le risposte. Usato per una varietà di scopi (cache, anonimato).

Pmt. Vedi →*photomultiplier tubes.*

Ppi. Acronimo di point per inch, pixel per pollice, è l'unità di misura della →*risoluzione* nelle immagini digitali.

Q

QoS vedi →*quality of service*

Quality of service. Indica i parametri usati per caratterizzare la qualità del servizio offerto dalla rete (ad esempio perdita di pacchetti, ritardo), o gli strumenti per ottenere una qualità di servizio desiderata.

Query. Termine inglese dell'informatica che sta per *interrogazione* (riferito a un database), indica, il processo attraverso cui si ottengono dati contenuti in una o più banche dati.

R

RAM. Acronimo dell'inglese *Random Access Memory*, memoria di tipo volatile, nel senso che mantiene le informazioni solo quando è sottoposta ad alimentazione. È inoltre il supporto di memoria su cui è possibile leggere e scrivere informazioni con un accesso "casuale", ovvero senza dover rispettare un determinato ordine, come ad esempio avviene per un nastro magnetico.

Raw. Formato di file naturale, registrato attraverso un particolare metodo di memorizzazione dei dati descrittori di un'immagine. Permette di non avere perdite di qualità della registrazione su un qualsiasi supporto rispetto ai segnali catturati dal →*sensore* e successivamente composti per →*interpolazione* dal processore d'immagine della fotocamera nelle sue tre componenti fondamentali →*RGB*.

Raster, tecnica utilizzata in computer grafica per descrivere un'immagine composta da →*pixel*

Record. Gruppo di campi correlati. Ciascun campo del record può contenere informazioni in formato diverso (es. caratteri o →*codice binario).*

Refreshing. Processo di riscrittura ciclica dei dati nella memoria

per evitare che questi vengano persi.

Retinatura. Tecnica fotografica utilizzata nell'industria grafica per simulare a stampa le variazioni chiaroscurali tipiche delle fotografie tradizionali.

Risoluzione. Quantità di elementi costitutivi di un'immagine per unità di misura. Maggiore è la risoluzione, migliore è la nitidezza dell'immagine.

Risoluzione ottica. Nelle immagini digitali, indica la densità dei →*pixel* rapportata ad una dimensione lineare di solito pollici (ppi, pixel per inch o dpi, dot per inch). Per alcuni dispositivi, la densità dei pixel è diversa nelle due dimensioni, come per gli scansionatori d'immagine, quindi occorre indicare sia la risoluzione orizzontale che quella verticale.

Risoluzione interpolata. →*Risoluzione* raggiunta attraverso l'utilizzo di →*algoritmi di* →*interpolazione d'immagine.*

RLE. Acronimo dell'inglese *Run-Length Encoding*, è storicamente il primo algoritmo di compressione per le immagini. È un algoritmo →*lossless,* ovvero permette di comprimere e decomprimere senza alcuna perdita di informazione.

ROM. Acronimo dell'inglese *Read Only Memory* (memoria di sola lettura), è un tipo di memoria che consente solamente l'operazione di lettura dei dati poiché il suo contenuto viene permanentemente definito in fase di costruzione della memoria stessa. È di tipo non volatile quindi il suo contenuto non viene perso se viene a mancare l'alimentazione.

RGB, modello di colori di tipo additivo che si basa sui tre colori Rosso (*Red*), Verde (*Green*) e Blu (*Blue*).

S

Sandwitch. sistema di protezione composto da teche di vetro, che permette di srotolare i documenti avvolti su se stessi senza subire pressioni dannose.

SCSI. Acronimo dell'inglese *Small Computer System Interface*, è un'interfaccia standard progettata per realizzare il trasferimento di dati fra diversi dispositivi interni di un computer collegati fra di loro tramite un →*bus.*

Scanner. →*Periferica* esterna in grado di acquisire in modalità ottica una superficie (fogli stampati, pagine di libri e riviste, fotografie, diapositive ecc.), di interpretarla come un insieme di →*pixel* e, quindi, di restituirne la copia sotto forma di immagine digitale.

Scansione. Vedi →*Scansire.*

Scansire. Dal verbo inglese *to Scan* (cioè separare in parti dette →*pixel*), è l'operazione che consente, attraverso l'uso di uno →*scanner*, la digitalizzazione in file di testi e immagini.

Scheda madre. In inglese motherboard, parte fondamentale di un moderno personal computer: raccoglie in sé tutta la circuiteria elettronica di interfaccia fra i vari componenti principali e fra questi e i →*bus* di espansione e le interfacce verso l'esterno.

Scheduling. operazione per il corretto ed efficiente funzionamento di uno o più calcolatori. Consente di eseguire più programmi concorrentemente e di migliorare le prestazioni.

Segnale analogico. Rappresentazione o trasformazione di una grandezza fisica tramite una sua analoga.

Segnale elettrico. Variazione di corrente o di tensione all'interno di un conduttore o in un punto di un circuito.

Segnale digitale. Segnale discreto, ovvero un segnale che può assumere solo un numero finito di valori rappresentati da numeri.

Sensore. Dispositivo che trasforma una grandezza fisica che si vuole misurare in un segnale di natura diversa (tipicamente elettrico) più facilmente misurabile o memorizzabile.

Server. Componente informatica che fornisce servizi ad altre componenti (tipicamente chiamate →*client*) attraverso una rete. Si noti che il termine server, così come pure il termine client, possono essere riferiti sia alla componente →*software* che alla componente →*hardware*.

Server side. Operazione che viene eseguita direttamente dal →*sistema remoto* a cui ci si sta collegando: l'utente riceve così direttamente il risultato dell'operazione, lasciando al server il costo – inteso come spazio e tempo di elaborazione – necessario per ottenerlo. Il sistema complementare è →*client side*.

SGML. Acronimo dell'inglese *Standard Generalized Markup Language*, è uno standard per la descrizione logica dei documenti, che utilizza un tipo di marcatura generica chiamata *marcatura descrittiva* che definisce la struttura logica dei documenti.

Sistema operativo. Programma responsabile del controllo e della gestione dei componenti hardware che costituiscono un computer e dei programmi che su di esso girano. Il sistema operativo mette anche a disposizione dei programmi una interfaccia →*software* per accedere alle risorse →*hardware* (dischi, memoria, I/O in generale) del sistema.

Sistema remoto. Sistema informatico presente su una rete diversa da quella a cui è collegato l'utente. L'accesso richiede il collegamento Internet, indipendentemente dalla distanza fisica.

Spazio colore. Combinazione di un modello di colore e di una appropriata funzione di mappatura di questo modello.

Software. Dall'inglese *soft* (morbido) e *ware* (manufatto, oggetto), indica un programma o un insieme di programmi in grado di funzionare su un elaboratore.

Start Tag. Nei linguaggi di marcatura indica l'inizio di un elemento. Uno start tag inizia con “<” (start) e finisce con “> “.

Stativo. Supporto, sostegno di vari strumenti, specialmente ottici.

T

Tag. Elemento sintattico con cui si marcano porzioni di un documento nei linguaggi di marcatura come HTML e XML.

TCP. Acronimo di *Transmission Control Protocol*, è il protocollo di livello di trasporto della suite di protocolli →*internet.*

Throughput. Riferito a un canale di comunicazione, rappresenta la sua capacità di trasmissione effettivamente utilizzata.

TIFF. Acronimo dell'inglese *Tagged Image File Format*, formato immagine di tipo →*raster* piuttosto diffuso e largamente utilizzato per lo scambio di immagini raster fra stampanti e scanner. Permette di specificare numerose indicazioni aggiuntive.

Tonalità. Colore "puro", ovvero caratterizzato da una singola lunghezza d'onda all'interno dello spettro visibile (o spettro ottico) della luce.

Transform coding. Tipo di compressione per dati “naturali”, come segnali audio e immagini fotografiche. La compressione è di

tipo →*lossy* e comporta una resa di minor qualità rispetto all'→*input* originale.

Truecolor. Modalità di memorizzazione delle informazioni dell'immagine nella memoria dei personal computer,conosciuta anche con l'espressione *milioni di colori*, nella quale il colore di ogni →*pixel* è codificato da tre o più →*byte*. In generale ogni byte esprime l'intensità di ciascuno dei tre canali →*RGB;* vi è anche un quarto byte che contiene dati sull'→*alpha channel* o viene semplicemente ignorato.

Twain. Standard di comunicazione tra computer e dispositivi di acquisizione di immagini come →*scanner* e →*fotocamere* digitali sotto forma di una →*API* per i →*sistemi operativi* Microsoft Windows e Apple Macintosh.

U

Unicode. sistema di codifica che assegna un numero univoco ad ogni carattere usato per la scrittura di testi, in maniera indipendente dalla lingua, dalla piattaforma informatica e dal programma utilizzati.

USB. Acronimo dell'inglese *Universal Serial* →*Bus*, standard di comunicazione seriale che consente di collegare diverse →*periferiche* a un computer.

W

W3C. Acronimo dell'inglese *World Wide Web Consortium*, organismo che ha lo scopo di definire standard e migliorare gli esistenti protocolli e linguaggi per il →*World wide web* e di aiutarlo a sviluppare tutte le sue potenzialità.

Wavelet. Si riferisce alla rappresentazione di un segnale mediante l'uso di una forma d'onda oscillante di lunghezza finita o a decadimento rapido (nota come wavelet madre).

Workflow. Gestione dei gruppi di lavoro collaborativi secondo un modello processuale che contempla più attività da svolgere per giungere a un obiettivo comune, grazie al sostegno di software specifici. Le attività possono essere svolte dai partecipanti o da applicazioni informatiche.

Workstation. computer da scrivania ad alte prestazioni, utilizzato da professionisti per il lavoro su disegno tecnico, ricerca scientifica, o produzioni audio/video.

World wide web. Spesso abbreviato in Web, è uno dei servizi di →internet, che mette a disposizione degli utenti uno spazio elettronico e digitale per la pubblicazione di contenuti multimediali, oltre che un mezzo per la distribuzione di software e la fornitura di servizi particolari sviluppati dagli stessi utenti. È stato creato da Tim Berners-Lee, mentre era ricercatore al CERN di Ginevra, e oggi gli standard su cui è basato, in continuo sviluppo, sono mantenuti dal *W3C.*

WWW. Acronimo di *World Wide Web*

X

XML. Acronimo dell'inglese *eXtensible Markup Language*, (linguaggio di marcatura estensibile) è un metalinguaggio creato e gestito dal *W3C.* È una semplificazione e adattamento dell'*SGML* e permette di definire la grammatica di diversi linguaggi specifici derivati.

19. Ringraziamenti

In breve: me, lamusa, l’arsenale, TC TdC, Floriana, Cesare.
Grazie.

Monica